

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:33:07 ; Search time 34.09 Seconds
(without alignments)
550.660 Million cell updates/sec

Title: US-09-508-418-2
Perfect score: 2844
Sequence: 1 MTTTPINHNHPIFHSSSS.....EGAYEVASEVTGELSRVAK 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2821	99.2	538	PROG_TOBAC	O24163 nicotiana t
2	2075	73.0	547	PROG_ARATH	P55826 arabidopsis
3	341	19.0	471	PROX_MYXXA	P56601 myxococcus
4	505	17.8	504	PROM_TOBAC	O24164 nicotiana t
5	465	17.1	470	PROX_BACSU	P32397 bacillus su
6	375.5	13.2	527	PROX_PROPR	O32434 propionibac
7	363.5	12.8	477	PROX_HUMAN	P50336 homo sapien
8	362.5	12.7	477	PROX_MOUSE	P51175 mus musculu
9	302.5	10.6	490	YAMF_SCHPO	Q10062 schizosacch
10	221	7.8	426	PROX_MCTU	O53230 mycobacteri
11	216.5	7.6	451	PROX_MYCLE	O50008 mycobacteri
12	189	6.6	539	PROX_YEAST	P40012 saccharomyc
13	141	5.0	448	ROFH_MYCTU	O53330 mycobacteri
14	139	4.9	570	CRTI_MARPS	O40406 narcissus p
15	134	4.7	522	ROF_ONCMY	P49223 oncorhynch
16	132.5	4.7	527	ROFA_BOVIN	P21338 bos taurus
17	131	4.6	566	CRTI_ARATH	O07356 arabidopsis
18	125.5	4.4	417	Y4AB_RHTSN	P55349 rhizobium s
19	125.5	4.4	472	CRTI_SYVW3	P29273 synchocyst
20	125.5	4.4	489	CRTI_CANAL	P11225 candida alb
21	123.5	4.3	526	ROFA_RAT	R11386 rattus norv
22	123	4.3	582	CRTI_CARAN	P80093 capsicum an
23	122	4.3	471	Y7B2_SYVW3	O55629 synchocyst
24	119.5	4.2	478	POO_MICRU	P40974 micrococcus
25	118	4.1	571	CRTI_MAIZE	P49086 zea mays (m
26	117.5	4.1	469	P49_STRLI	P06108 streptomyce
27	116	4.0	583	CRTI_LYCES	P28554 lycopersico
28	114.5	4.0	1106	IRA7_RAT	O63258 rattus norv
29	114	4.0	531	AGHC_ACHFU	P35993 achentina fu
30	112.5	4.0	570	CRTI_SOYBN	P28553 glycine max
31	111	3.9	508	CRTI_STRSE	P54971 streptomyce
32	109.5	3.9	527	ROFA_HUMAN	P21397 homo sapien
33	109.5	3.9	908	SRCA_RABIT	P13666 oryctolagus

RESULT ID	PROG_TOBAC	STANDARD:	PRT:	548 AA.	ALIGNMENTS
AC	O24163;				
AD	15-DEC-1998 (Rel. 37, Created)				
AE	15-DEC-1998 (Rel. 37, Last sequence update)				
AF	15-JUL-1999 (Rel. 38, Last annotation update)				
AG	PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO I)				
AH	(PROTOPORPHYRINOGEN IX OXIDASE ISOZYME I) (PPX I).				
AI	PPXI.				
AJ	Nicotiana tabacum (Common tobacco).				
AK	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;				
AL	Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;				
AM	Solanales; Solanaceae; Nicotiana.				
AN	NCBI_TaxID=4097;				
AO	[1]				
AP	SEQUENCE FROM N.A.				
AQ	STRAIN=CV. SRL;				
AR	MEDLINE=97385200; Pubmed=9238074;				
AS	Lermontova I., Kruse E., Mock H.-P., Grimm B.;				
AT	Cloning and characterization of a plastidial and a mitochondrial				
AV	isoform of tobacco protoporphyrinogen IX oxidase. #;				
AW	Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997).				
AX	RT				
AY	1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN				
AZ	IX TO FORM PROTOPORPHYRIN IX.				
BA	1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-				
BB	IX + H(2)O(2).				
BC	1- COFACTOR: FAD (BY SIMILARITY).				
BD	1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS. AND				
BE	CHLOROPHYLL SYNTHESIS. CONTAINS ONE FAD PER HOMODIMER (BY				
BF	1- SUBUNIT: HOMODIMER.				
BG	1- SIMILARITY:				
BH	1- SUBCELLULAR LOCATION: CHLOROPLAST.				
BI	1- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREATURE LEAVES.				
BJ	1- DECREASED EXPRESSION IN OLDEST LEAVES. EXPRESSED AT VERY LOW LEVEL				
BK	IN ROOTS.				
BL	1- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL				
BM	EXPRESSION IN THE DARK PERIOD.				
BN	1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.				
BO	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION				
BP	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
BQ	the European Bioinformatics Institute. There are no restrictions on its				
BR	use by non-profit institutions as long as its content is in no way				
BS	modified and this statement is not removed. Usage by and for commercial				
BT	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
BU	or send an email to license@isb-sib.ch).				
BV	EMBL: Y13465; CAA73865.1;				
BW	porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;				
BX	Transit peptide; Flavoprotein; FAD; Chloroplast;				
BY	Chlorophyll biosynthesis.				
BZ	TRANSIT				
CA	CHLOROPLAST (POTENTIAL).				
CB	PROTOPORPHYRINOGEN OXIDASE.				
CC	FAD (POTENTIAL).				
CD	NP_BIND				
CE	SEQUENCE				
CF	548 AA;				
CG	59230 MW;				
CH	66892878FBBA3E30 CR664;				

Query Match 99.2% Score 2821; DB 1; Length 548;
 Best Local Similarity 99.3% Pred. No. 14e-205;
 Matches 544; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	1	MTTPTIANHPNIFTHOSSSPLAFINRTSIFPSSISKRNYSVNCMGRPCSVAKDYVTP	60
DB	1	MTTPTIANHPNIFTHOSSSPLAFINRTSIFPSSISKRNYSVNCMGRPCSVAKDYVTP	60
QY	61	SSAVDGGPAAELDCVYVYAGISGLCAQVMSANYPMLWYTKARDPAGNITVVERDGYLW	120
DB	61	SSAVDGGPAAELDCVYVYAGISGLCAQVMSANYPMLWYTKARDPAGNITVVERDGYLW	120
QY	121	EEGPNPSPDPMETMAVVDGKDDVLDDPPNAPRFLVLMKGLRBPVSKLIDLDPEFDLMS	180
DB	121	EEGPNPSPDPMETMAVVDGKDDVLDDPPNAPRFLVLMKGLRBPVSKLIDLDPEFDLMS	180
QY	181	IPGKLRAGFGPIGLRPSPPGHEBSVQFVRRNLGGEFVERLLEPPSCGVYVGDPSKLSMK	240
DB	181	IPGKLRAGFGPIGLRPSPPGHEBSVQFVRRNLGGEFVERLLEPPSCGVYVGDPSKLSMK	240
QY	241	AAFRWVWKLBEETGSIITGGTFKAIKRSSTPRAPRDPRLPKKGGTVSSFRKGLRMLPDA	300
DB	241	AAFRWVWKLBEETGSIITGGTFKAIKRSSTPRAPRDPRLPKKGGTVSSFRKGLRMLPDA	300
QY	301	ISARLGSKLLKMSKLSITSEKGGYHLTETPEGVVSHOSRSIYMTVPYVVAANILRPL	360
DB	301	ISARLGSKLLKMSKLSITSEKGGYHLTETPEGVVSHOSRSIYMTVPYVVAANILRPL	360
QY	361	SVAADADLSLNFYYPVGAATITTPQPAIRDELVLVDELKGFQDLPRFGVETLGTITSS	420
DB	361	SVAADADLSLNFYYPVGAATITTPQPAIRDELVLVDELKGFQDLPRFGVETLGTITSS	420
QY	421	SLPFRNAPKGRVLLNVTGGAKNPELSTRESOLVYVDRDRKMLIKPKKODDPLVYVGR	480
DB	421	SLPFRNAPKGRVLLNVTGGAKNPELSTRESOLVYVDRDRKMLIKPKKODDPLVYVGR	480
QY	481	VMPQAIPOPLVGHLDLSTFAKAMNDNGLEGLFLGNNVYSGVALRCRVEGAYEVAASEVTG	540
DB	481	VMPQAIPOPLVGHLDLSTFAKAMNDNGLEGLFLGNNVYSGVALRCRVEGAYEVAASEVTG	540
QY	541	FLSRVAYK 548	
DB	541	FLSRVAYK 548	

RESULT 2
 PROC_ARATH STANDARD; PRT; 537 AA.

AC P55826;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, last sequence update)
 DT 01-OCT-2000 (rel. 40, last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO).
 GN PPOX OR AT4G01990 OR T15B16.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RN MEDLINE=9716707; PubMed=8982084;
 RN Naftis S., Tanaka R., Ito T., Okada K., Taketani S., Inokuchi H.,
 RR "Molecular cloning and characterization of a cDNA that encodes
 RR protoporphyrinogen oxidase of Arabidopsis thaliana.";
 RL Gene 182:169-175(1996).
 RP [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=CV. COLUMBIA;
 RP MEDLINE=20083488; PubMed=10617198;
 RP "K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,

RA Pohl T., Duesterhoeft A., Stiekema W., Ertlan K.-D., Terryn N.,
 RA Harris B., Ansoerger W., Brandt P., Griwell L., Rieger M.,
 RA Melchelsgärtner M., de Simone V., Obermaler B., Macho R., Mueller M.,
 RA Kreis M., Delseny M., Puligdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohnsbein J., Zimmermann W., Medler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandennusche F.,
 RA Braeken M., Meljens I., Voet M., Bastiaens I., Aert R., Defoore F.,
 RA Weiztenegger T., Bothe G., Kampsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., Van Staveren M., Dirxer W.,
 RA Mooljman P., Klein Lankehorst R., Rose M., Hauf J., Koetler P.,
 RA Berleiser S., Hempel S., Feldpausch M., Lambers S., Van den Daele H.,
 RA De Keyser A., Buysschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Tennard N., Weclay K., Hayes R.,
 RA Petrett A., Rajandream M.-A., Lyne M., Benes V., Reclmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T. H.,
 RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argiriou A., Vitale D., Lignori R., Piravandi E.,
 RA Massenot O., Outgley F., Clabaud G., Mueandlein A., Felber R.,
 RA Schabdi S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedot F., Cooke R., Berger C., Montfort A., Casacuberta E.,
 RA Gidbons T., Weber N., Vandenbol M., Barygues M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Hejlsman D., Schwarz S., Scholler P., Heber S., Frances P., Belke C.,
 RA Fajman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Beyan M., Nilsson R.K., de la Bastide M., Hedermann K.,
 RA Parnell L., Dehnia N., Guoj L., Schutz K., Huang E., Spliegel L.,
 RA Senon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stonking T., Kallio J., Graves T., Harmon G., Edwards J.,
 RA Lacroix P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mix P., Bentley D., Fulton B., Miller N., Grego T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spleth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali Z., Bergthoff A., Jones K., Dione K., Cotton M., Joshi C.,
 RA Anonoulu B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vill D., Sheher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Grant S., Shobhy N., Hasegawa A., Hammed A., Lodi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.,
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RA thaliana."
 RL Nature 402:769-777(1999).
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 CC IX TO FORM PROTOPORPHYRIN IX.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2). (BY SIMILARITY).
 CC -1- COFACTOR: FAD (BY SIMILARITY).
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE LEAVES AND AT
 CC LOW LEVELS IN THE ROOTS AND FLORAL BUDS.
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.embnet.org/licenses/embnet.html>)
 CC or send an email to license@ebi.ac.uk.
 CC
 DR EMBL: D83139; BA011820.1;
 DR EMBL: AL161492; CAB77739.1;
 KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
 KW Flavoprotein; FAD; Chloroplast; Transferrin;
 FT TRANSIT 1 537 CHLOROPLAST (POTENTIAL).
 FT CHAIN 2 537 PROTOPORPHYRINOGEN OXIDASE.
 FT NP_BIND 63 68 FAD (POTENTIAL).
 FT SEQUENCE 537 AA; 57695 MW; C125B0EEED13187D5 CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:00:01 ; Search time 28.95 Seconds

(Without alignments)
381.322 Million cell updates/sec

Title: US-09-508-418-2
2844

Sequence: 1 MTTTPINRRIRFHSSSS.....EGAYEAVSEVTGLSRRAVK 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144636 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCNUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2113	74.3	539	2 US-08-808-931-16	Sequence 16, App1
2	2113	74.3	539	3 US-08-808-323-16	Sequence 16, App1
3	2113	74.3	539	3 US-09-050-603A-16	Sequence 16, App1
4	2113	74.3	539	3 US-09-102-420B-16	Sequence 16, App1
5	2075	73.0	537	2 US-08-472-028A-2	Sequence 2, App1
6	2075	73.0	537	1 US-08-808-931-2	Sequence 2, App1
7	2075	73.0	537	3 US-08-808-323-2	Sequence 2, App1
8	2075	73.0	537	3 US-09-050-603A-2	Sequence 2, App1
9	2075	73.0	537	3 US-09-102-420B-2	Sequence 2, App1
10	2075	72.3	543	4 US-09-071-296-2	Sequence 2, App1
11	2055	72.3	543	4 US-08-808-931-12	Sequence 12, App1
12	2055	72.3	543	3 US-08-808-323-12	Sequence 12, App1
13	2055	72.3	543	3 US-09-050-603A-12	Sequence 12, App1
14	2055	72.3	543	3 US-09-102-420B-12	Sequence 12, App1
15	2006.5	70.6	536	2 US-08-808-931-20	Sequence 20, App1
16	2006.5	70.6	536	2 US-08-808-323-20	Sequence 20, App1
17	2006.5	70.6	536	3 US-09-050-603A-20	Sequence 20, App1
18	2006.5	70.6	536	3 US-09-102-420B-20	Sequence 20, App1
19	1987	69.9	528	2 US-08-808-931-10	Sequence 10, App1
20	1987	69.9	528	2 US-08-808-323-10	Sequence 10, App1
21	1987	69.9	528	3 US-09-050-603A-10	Sequence 10, App1
22	1987	69.9	528	3 US-09-102-420B-10	Sequence 10, App1
23	1978	69.5	481	1 US-08-472-028A-6	Sequence 6, App1
24	1978	69.5	481	2 US-08-808-931-6	Sequence 6, App1
25	1978	69.5	481	3 US-08-808-323-6	Sequence 6, App1
26	1978	69.5	481	3 US-09-050-603A-6	Sequence 6, App1
27	1978	69.5	481	3 US-09-102-420B-6	Sequence 6, App1

Result No.	Score	Query Match %	Length	DB ID	Description
28	1978	69.5	483	4 US-09-071-296-6	Sequence 6, App1
29	1933.5	68.0	560	2 US-08-808-931-18	Sequence 18, App1
30	1933.5	68.0	560	3 US-08-808-323-18	Sequence 18, App1
31	1933.5	68.0	560	3 US-09-050-603A-18	Sequence 18, App1
32	1933.5	68.0	560	3 US-09-102-420B-18	Sequence 18, App1
33	1823	64.1	440	2 US-08-808-931-24	Sequence 24, App1
34	1823	64.1	440	3 US-08-808-323-24	Sequence 24, App1
35	1823	64.1	440	3 US-09-050-603A-24	Sequence 24, App1
36	1823	64.1	440	3 US-09-102-420B-24	Sequence 24, App1
37	1261	44.3	312	2 US-08-808-931-22	Sequence 22, App1
38	1261	44.3	312	3 US-08-808-323-22	Sequence 22, App1
39	1261	44.3	312	3 US-09-050-603A-22	Sequence 22, App1
40	1261	44.3	312	3 US-09-102-420B-22	Sequence 22, App1
41	490.5	17.2	508	1 US-08-472-028A-4	Sequence 4, App1
42	490.5	17.2	508	2 US-08-808-931-4	Sequence 4, App1
43	490.5	17.2	508	3 US-08-808-323-4	Sequence 4, App1
44	490.5	17.2	508	3 US-09-050-603A-4	Sequence 4, App1
45	490.5	17.2	508	3 US-09-102-420B-4	Sequence 4, App1

ALIGNMENTS

RESULT 1:
US-08-808-931-16
Sequence 16, Application US/08808931
Patent No. 5939602
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Pater, Sharon
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5939602artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,931
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38, 241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8687
TELEFAX: (919) 341-8689
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid

STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-808-931-16

Query Match
Best Local Similarity 74.3%, Score 2113; DB 2; Length 539;
Matches 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;

47 WRRCGVAKDYTPSSAAVDGPAELDCVYAGISGLCIAQVNSANY----PMLWTEA 102
34 FKLRCGLAAGPPISSKIDGSSVADCVYGGISGLCIAQALATLHRVANSVIVTER 93
103 RDRAGNITTVBERDGLYMEBGNFQSPDPMITMAVDSGLKDDLVLDGPNAPRFVLMK 162
94 RDRVGNITTVBERDGLYMEBGNFQSPDPMITMAVDSGLKDDLVLDGPNAPRFVLMK 153
163 LRPVSKLTDLPFDDMSIPGKLRAGFGPIGLRSPSPGHEBSVQFVRRMLGGEVFERLI 222
154 LRPVSKPTDLPFDDMSIPGKLRAGFGPIGLRSPSPGHEBSVQFVRRMLGGEVFERLI 213
223 EPPCGSVYVADPDKLSMKAAFGKWKLEBETGSGIIGGTFAIKERSSTPKAPRDPRLPK 282
214 EPPCGSVYVADPDKLSMKAAFGKWKLEBETGSGIIGGTFAIKERSSTPKAPRDPRLPK 273
283 KQGVVSPFRKGLRMLPDAISARLGSKLRKLSMTKLSITKSEKGYHLTYETPEGVVSIQSR 342
274 KQGVVSPFRKGLRMLPDAISARLGSVNVKLSMTKLSITKLSNGVNLTFEPEGVVSIQSR 333
343 SIYWTVPYSVAANIPLPLSVAADALSNFYYPVAVTISYPOEALIDELVDGELKFG 402
334 SVYWTIPSHVANSMLHPLSAAADALSOFYYPVAVTISYPOEALIDELVDGELKFG 393
403 QLHPRTQGVETLGTIYSSSLFPNRAKGRVLLNLTIGGAKNPELSTESQVLEVVDRDL 462
394 QLHPRSQGIETLGTIYSSSLFPNRAKGRVLLNLTIGGAKNPELSTESQVLEVVDRDL 453
463 RKMLIRKADDPVYGVVRWPAIQPLVGHLDLSTAKAAMNDNGLEGLLGGVYVSGV 522
454 RKMLINNAKDPVYGVVRWPAIQPLVGHLDLSDAKMALRDSGFGHLFDSGNVYSGV 513
523 ALGRCEGAYEVAEVTGFLSRVAYK 548
514 ALGRCEGAYEVAEYKFLSOYAYK 539

RESULT 2
US-08-808-323-16
Sequence 16, Application US/08808323

Patent No. 6018105
GENERAL INFORMATION:
APPLICANT: Johnson, Marie
APPLICANT: Volirath, Sandra
APPLICANT: Ward, Eric
TITLE OF INVENTION: Promoters from plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 6018105, Atlas Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005

COMPUTER READABLE FORM:
MEDIA TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
APPLICATION DATA:
APPLICATION NUMBER: US/08/808.323

APPLICANT: Johnson, Marie
APPLICANT: Volirath, Sandra
APPLICANT: Ward, Eric
TITLE OF INVENTION: Promoters from plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 6018105, Atlas Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-808-323-16

Query Match
Best Local Similarity 74.3%, Score 2113; DB 3; Length 539;
Matches 461; Conservative 46; Mismatches 55; Indels 4; Gaps 1;

47 WRRCGVAKDYTPSSAAVDGPAELDCVYAGISGLCIAQVNSANY----PMLWTEA 102
34 FKLRCGLAAGPPISSKIDGSSVADCVYGGISGLCIAQALATLHRVANSVIVTER 93
103 RDRAGNITTVBERDGLYMEBGNFQSPDPMITMAVDSGLKDDLVLDGPNAPRFVLMK 162
94 RDRVGNITTVBERDGLYMEBGNFQSPDPMITMAVDSGLKDDLVLDGPNAPRFVLMK 153
163 LRPVSKLTDLPFDDMSIPGKLRAGFGPIGLRSPSPGHEBSVQFVRRMLGGEVFERLI 222
154 LRPVSKPTDLPFDDMSIPGKLRAGFGPIGLRSPSPGHEBSVQFVRRMLGGEVFERLI 213
223 EPPCGSVYVADPDKLSMKAAFGKWKLEBETGSGIIGGTFAIKERSSTPKAPRDPRLPK 282
214 EPPCGSVYVADPDKLSMKAAFGKWKLEBETGSGIIGGTFAIKERSSTPKAPRDPRLPK 273
283 KQGVVSPFRKGLRMLPDAISARLGSKLRKLSMTKLSITKSEKGYHLTYETPEGVVSIQSR 342
274 KQGVVSPFRKGLRMLPDAISARLGSVNVKLSMTKLSITKLSNGVNLTFEPEGVVSIQSR 333
343 SIYWTVPYSVAANIPLPLSVAADALSNFYYPVAVTISYPOEALIDELVDGELKFG 402
334 SVYWTIPSHVANSMLHPLSAAADALSOFYYPVAVTISYPOEALIDELVDGELKFG 393
403 QLHPRTQGVETLGTIYSSSLFPNRAKGRVLLNLTIGGAKNPELSTESQVLEVVDRDL 462
394 QLHPRSQGIETLGTIYSSSLFPNRAKGRVLLNLTIGGAKNPELSTESQVLEVVDRDL 453
463 RKMLIRKADDPVYGVVRWPAIQPLVGHLDLSTAKAAMNDNGLEGLLGGVYVSGV 522
454 RKMLINNAKDPVYGVVRWPAIQPLVGHLDLSDAKMALRDSGFGHLFDSGNVYSGV 513
523 ALGRCEGAYEVAEVTGFLSRVAYK 548
514 ALGRCEGAYEVAEYKFLSOYAYK 539

RESULT 3
US-09-050-603A-16
Sequence 16, Application US/09050603A
Patent No. 6023012
GENERAL INFORMATION:

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:00:01 ; Search time 28.95 Seconds

(Without alignments)
381,322 Million cell updates/sec

Title: US-09-508-418-2

Perfect score: 28.44
Sequence: 1 MTRPPIANHPNPIFHQSSSS.....EGAYEVASEVTLGFLSRVYAK 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 segs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:**
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:**
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:**
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:**
- 5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:**
- 6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2113	74.3	539	2	US-08-808-931-16
2	2113	74.3	539	3	US-08-808-323-16
3	2113	74.3	539	3	US-09-050-603A-16
4	2113	74.3	539	3	US-09-102-420B-16
5	2075	73.0	537	1	US-08-472-028A-2
6	2075	73.0	537	1	US-08-808-931-2
7	2075	73.0	537	3	US-08-808-323-2
8	2075	73.0	537	3	US-09-050-603A-2
9	2075	73.0	537	3	US-09-102-420B-2
10	2055	72.3	543	2	US-08-808-931-12
11	2055	72.3	543	2	US-08-808-323-12
12	2055	72.3	543	3	US-09-050-603A-12
13	2055	72.3	543	3	US-09-102-420B-12
14	2006.5	70.6	536	2	US-08-808-931-20
15	2006.5	70.6	536	2	US-08-808-323-20
16	2006.5	70.6	536	3	US-09-050-603A-20
17	2006.5	70.6	536	3	US-09-102-420B-20
18	1987	69.9	528	2	US-08-808-931-10
19	1987	69.9	528	2	US-08-808-323-10
20	1987	69.9	528	3	US-09-050-603A-10
21	1987	69.9	528	3	US-09-102-420B-10
22	1987	69.9	528	3	US-08-808-931-10
23	1978	69.5	481	1	US-08-472-028A-6
24	1978	69.5	481	2	US-08-808-931-6
25	1978	69.5	481	3	US-08-808-323-6
26	1978	69.5	481	3	US-09-050-603A-6
27	1978	69.5	481	3	US-09-102-420B-6

Result No.	Score	Query Match	Length	ID	Description
28	1978	69.5	483	4	US-09-071-296-6
29	1933.5	68.0	560	2	US-08-808-931-18
30	1933.5	68.0	560	3	US-08-808-323-18
31	1933.5	68.0	560	3	US-09-050-603A-18
32	1933.5	68.0	560	3	US-09-102-420B-18
33	1823	64.1	440	2	US-08-808-931-24
34	1823	64.1	440	3	US-08-808-323-24
35	1823	64.1	440	3	US-09-050-603A-24
36	1823	64.1	440	3	US-09-102-420B-24
37	1261	44.3	312	2	US-08-808-931-22
38	1261	44.3	312	2	US-08-808-323-22
39	1261	44.3	312	3	US-09-050-603A-22
40	1261	44.3	312	3	US-09-102-420B-22
41	490.5	17.2	508	1	US-08-472-028A-4
42	490.5	17.2	508	2	US-08-808-931-4
43	490.5	17.2	508	3	US-08-808-323-4
44	490.5	17.2	508	3	US-09-050-603A-4
45	490.5	17.2	508	3	US-09-102-420B-4

ALIGNMENTS

```

RESULT 1
US-08-808-931-16
; Sequence 16, Application US/08808931
; Patent No. 59396602
; GENERAL INFORMATION:
; APPLICANT: Voltrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potler, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding plant
; TITLE OF INVENTION: Protocopyrinogen oxidase and inhibitor-Resistant Mutants
; TITLE OF INVENTION: Therocof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5939602artlis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 amino acids
; TYPE: amino acid

```

STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-808-931-16

Query Match 74.3% Score 2113; DB 2; Length 539;
Best Local Similarity 79.2%; Pred. No. 9.4e-202;
Matches 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;

47 WTRVCYVAKDYTPSSAVDGGPAEELDCVTVGAGISGLCIAQVMSANY----PMLMATEA 102
34 FKLRCSLAEBEPTLSSSKIDGESSIADCVTVGGGIGSGLCIAQALATKRHDVAASNVIVTEA 93
103 RDRAGNITTVREDGYLMEEGPNSFQPSDPIITMAVDSGLKDDLVLGDPNAPRFVLMGK 153
94 RDRVGNITTVREDGYLMEEGPNSFQPSDPIITMAVDSGLKDDLVLGDPNAPRFVLMGK 153
163 LRPVSKLTLDPFDLMSIPGKLRAGFPIGLRPPGHESEVEQFVRRNLGGEVFERLI 222
154 LRPVSKLPDLPFDLMSIPGKLRAGFPIGLRPPGHESEVEEYRRLNGAIVEFERFI 213
223 EPPCGVYVGDPSKLSMKAAFGKWKLEETGSGIIGTFKAIKERSSTPKAPRDRPKP 282
214 EPPCGVYVGDPSKLSMKAAFGKWKLEETGSGIIGTFKTIQERNKTRPKRDRPKP 273
283 KQGTVGSFRKGLRMLPDAISARLGSKIKLSWKLSSTTKSEKGGYHLTYEPEGVYVLSQSR 342
274 KQGTVGSFRKGLRMLPDAISARLGSKIKLSWKLSSTTKKLGNGYMLTFEPEBGMVLSQSR 333
343 SIYMTVSVYVANSILRPLSVAADALSNFYYPVAVTISYPOEAIKIRDERLVDEGLKFG 402
334 SVYMTIPSHVANSNLHPLSAAADALDLSQFYYPVAVTISYPOEAIKIRDELDLIDGELKFG 393
403 QLHPRQGVETLGTIYSSSLFPNRPAPKGRVLLNLTGKANKPILSKTESQLVYVVDRL 462
394 QLHPRQGVETLGTIYSSSLFPNRPAPKGRVLLNLTGKANKPILSKTESQLVYVVDRL 453
463 RKMILKRAKADPLVYGVVRWVPOAIPQFLVGHLDPTLSTAKAAMNDNGLEGLFVYVYVSGV 522
454 RKMILNPNAKDPLVYGVVRWVPOAIPQFLVGHLDLDSAKAMLRDSDGFRHGLFVYVYVSGV 513
523 ALGRVCVGAIVEADEVYGLFVYVYVYK 548
514 ALGRVCVGAIVEADEVYGLFVYVYVYK 539
RESULT 2
US-08-808-323-16
; Sequence 16, Application US/08808323
; Patent No. 6018105
GENERAL INFORMATION:
APPLICANT: Johnson, Maile
APPLICANT: Voltrath, Sandra
APPLICANT: Ward, Eric
TITLE OF INVENTION: Promoters from Plant
TITLE OF INVENTION: Protocophylinogen Oxidase Genes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6018105artls Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808.323
FILING DATE:

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-808-323-16
Query Match 74.3% Score 2113; DB 3; Length 539;
Best Local Similarity 79.2%; Pred. No. 9.4e-202;
Matches 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;
47 WTRVCYVAKDYTPSSAVDGGPAEELDCVTVGAGISGLCIAQVMSANY----PMLMATEA 102
34 FKLRCSLAEBEPTLSSSKIDGESSIADCVTVGGGIGSGLCIAQALATKRHDVAASNVIVTEA 93
103 RDRAGNITTVREDGYLMEEGPNSFQPSDPIITMAVDSGLKDDLVLGDPNAPRFVLMGK 162
94 RDRVGNITTVREDGYLMEEGPNSFQPSDPIITMAVDSGLKDDLVLGDPNAPRFVLMGK 153
163 LRPVSKLTLDPFDLMSIPGKLRAGFPIGLRPPGHESEVEQFVRRNLGGEVFERLI 222
154 LRPVSKLPDLPFDLMSIPGKLRAGFPIGLRPPGHESEVEEYRRLNGAIVEFERFI 213
223 EPPCGVYVGDPSKLSMKAAFGKWKLEETGSGIIGTFKAIKERSSTPKAPRDRPKP 282
214 EPPCGVYVGDPSKLSMKAAFGKWKLEETGSGIIGTFKTIQERNKTRPKRDRPKP 273
283 KQGTVGSFRKGLRMLPDAISARLGSKIKLSWKLSSTTKSEKGGYHLTYEPEGVYVLSQSR 342
274 KQGTVGSFRKGLRMLPDAISARLGSKIKLSWKLSSTTKKLGNGYMLTFEPEBGMVLSQSR 333
343 SIYMTVSVYVANSILRPLSVAADALSNFYYPVAVTISYPOEAIKIRDERLVDEGLKFG 402
334 SVYMTIPSHVANSNLHPLSAAADALDLSQFYYPVAVTISYPOEAIKIRDELDLIDGELKFG 393
403 QLHPRQGVETLGTIYSSSLFPNRPAPKGRVLLNLTGKANKPILSKTESQLVYVVDRL 462
394 QLHPRQGVETLGTIYSSSLFPNRPAPKGRVLLNLTGKANKPILSKTESQLVYVVDRL 453
463 RKMILKRAKADPLVYGVVRWVPOAIPQFLVGHLDPTLSTAKAAMNDNGLEGLFVYVYVSGV 522
454 RKMILNPNAKDPLVYGVVRWVPOAIPQFLVGHLDLDSAKAMLRDSDGFRHGLFVYVYVSGV 513
523 ALGRVCVGAIVEADEVYGLFVYVYVYK 548
514 ALGRVCVGAIVEADEVYGLFVYVYVYK 539
RESULT 3
US-09-050-603A-16
; Sequence 16, Application US/09050603A
; Patent No. 6023012
GENERAL INFORMATION:

```

? APPLICANT: Volrath, Sandra
? APPLICANT: Johnson, Marie
? APPLICANT: Potter, Sharon
? APPLICANT: Ward, Eric
? APPLICANT: Helfetz, Peter
? TITLE OF INVENTION: DNA Molecules Encoding Plant
? TITLE OF INVENTION: Photoporphyrinogen Oxidase
? NUMBER OF SEQUENCES: 37
? CORRESPONDENCE ADDRESS:
? ADDRESS: No. 6023012artis Corporation
? STREET: 3054 Cornwallis Road
? CITY: Research Triangle Park
? STATE: NC
? COUNTRY: USA
? ZIP: 27709
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/050,603A
? FILING DATE: 30-MAR-1998
? CLASSIFICATION: 800
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/808,931
? FILING DATE: 28-FEB-1997
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/012,705
? FILING DATE: 28-FEB-1996
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/013,612
? FILING DATE: 28-FEB-1996
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/020,003
? FILING DATE: 21-JUN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Meigs, J. Timothy
? REGISTRATION NUMBER: 38,241
? REFERENCE/DOCKET NUMBER: CGC 1847
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (919) 541-8587
? TELEFAX: (919) 541-8689
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 539 amino acids
? TYPE: amino acid
? STRANDEDNESS: not relevant
? TOPOLOGY: not relevant
? MOLECULE TYPE: protein
? US-09-050-603A-16

```

```

Query Match 74.3%; Score 2113; DB 3; Length 539;
Best Local Similarity 79.2%; Pred. No. 9,4e-202;
Matches 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;

```

```

QY 47 WTRCSVAKDVTYSSAVDGGPARALDQVYVAGAGISGLCIAQVMSANY-----PMLMVTGA 102
Db 34 FKLRSGLAEGPRTSSKIDGSSSTADCVIVGGISGLCIAQALATKTRDVAASNYVTEA 93
QY 103 RDRAGNITTVRSDGLYMEEGNSFQSPDPMLTMAVDDGLDLDIVLGGPNAPRYVLMKGG 162
Db 94 RDRVGNITTVRSDGLYMEEGNSFQSPDPMLTMAVDDGLDLDIVLGGPNAPRYVLMKGG 153
QY 153 LRPVSKLTDLPFDLMSIPKGLRAGFGPRLGRSPSPGHEBSVQDFVFRNLDGGEVFERLI 222
Db 154 LRPVSKLTDLPFDLMSIPKGLRAGFGAIGIRPPRPGEBSVDFVFRNLDGGEVFERFI 213
QY 223 EPCFSGVAVGSPSKLMSKAARGKVMKLEEMTGGSTITGGFKAKRKSSTPPKAPRDLPKP 282
Db 214 EPCFSGVAVGSPSKLMSKAARGKVMKLEEMTGGSTITGGFKAKRKSSTPPKAPRDLPKP 273

```

```

QY 283 KQGVGSPFRKGLRMLPDALISARLGSKIKLSWKSITSEKGGYHLLTETPEGVSIQSR 342
Db 274 KQGVGSPFRKGLTMLPEALANSLSGNVSKLSWKSITSLGNGYVLTETPEGVSIQSR 333
QY 343 STVMTVPSYVASNILRPLSVAAAADALSNFYPPYPPVAVVTISYPOEARIRDERLVDLKGFG 402
Db 334 SVVMTTPSHVASNLLHPISAAAADALSGFYPPVAVSVYVKEARIRRECLIDDELKGF 393
QY 403 QHPRVQGVETLGTYSSTLPPNRPKGRVILLNVTGAKNPELISKTESQVVEVYDRDL 462
Db 394 QHPRVQGVETLGTYSSTLPPNRPKGRVILLNVTGAKNPELISKTESQVVEVYDRDL 453
QY 463 RKMLIKPKAODELVYGVVWMPQAIPOFLVGHDDTLSTAKAAMNDNGLEGLFLGGNYSV 522
Db 454 RKMLINPNKADLVYGVVWMPKAIPOFLVGHDDLLDSAKKALRDSGRHGLFLGGNYSV 513
QY 523 ALGRCEGAVEYASVETGFLSRAYK 548
Db 514 ALGRCEGAVEYAAAEVKKFLSOYAYK 539

```

```

RESULT 4
US-09-102-420B-16
? Sequence 16, Application US/09102420B
? Patent No. 6084155
? GENERAL INFORMATION:
? APPLICANT: Volrath, Sandra
? APPLICANT: Johnson, Marie
? APPLICANT: Ward, Eric
? APPLICANT: Helfetz, Peter
? TITLE OF INVENTION: HERBICIDE-TOLERANT PHOTOPORPHYRINOGEN
? TITLE OF INVENTION: OXIDASE ("PROTOX")
? NUMBER OF SEQUENCES: 43
? CORRESPONDENCE ADDRESS:
? ADDRESS: No. 6084155artis Corporation
? STREET: 3054 Cornwallis Road
? CITY: Research Triangle Park
? STATE: NC
? COUNTRY: USA
? ZIP: 27709
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/102,420B
? FILING DATE: 22-JUN-1998
? CLASSIFICATION: 800
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 09/059,164
? FILING DATE: 13-APR-1998
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 09/050,603
? FILING DATE: 30-MAR-1998
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/126,430
? FILING DATE: 11-MAR-1998
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/808,931
? FILING DATE: 28-FEB-1997
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/012,705
? FILING DATE: 28-FEB-1996
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/013,612
? FILING DATE: 28-FEB-1996
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/020,003
? FILING DATE: 21-JUN-1996
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/472,028
? FILING DATE: 06-JUN-1995

```



```

: APPLICANT: Volrath, Sandra
: APPLICANT: Johnson, Marie
: APPLICANT: Potter, Sharon
: APPLICANT: Ward, Eric
: APPLICANT: Helfetz, Peter
: TITLE OF INVENTION: DNA Molecules Encoding Plant
: TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
: TITLE OF INVENTION: thereof
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESS: No. 593962artis Corporation
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591-9005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/808,931
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/012,705
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/013,612
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/020,003
: FILING DATE: 21-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: CGC 1847
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 537 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-808-931-2

```

```

Query Match 73.0%; Score 2075; DB 2; Length 537;
Best Local Similarity 75.4%; Pred. No. 5,7e-198;
Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;
OY 22 LAFNRFSFIEFSSISKSNVNGW---RRGCVADYTPSSAVDGGPAEL--DCVI 76
Db | : | | | | | : | | | | | : | | | | | | | | | | | | | | | |
77 VGAGISGICIAQVYANYSAN-----PILMVTVEARDRAGNITVEYRDGYLMEEGPNSFQSPD 132
Db | : | | | | : | | | | | | | | | | | | | | | | | | | | | |
62 VGGISGICIAQALATKHPDAAPMLVITEAKDRGGINITREBNGPFLMEBEGPNSFQSSDP 121
OY 133 MLTMAVDCGLKDDVLVLDGPNAPRVVLMKGLRPVPSKLTLDLPPFDLMSIPGKLRAFGPI 192
Db | : | | | | : | | | | | | | | | | | | | | | | | | | | | |
122 MLTMAVDSGLKDDVLVLDGPNAPRVVLMKGLRPVPSKLTLDLPPFDLMSIPGKLRAFGPI 181
OY 193 GIRSPPGHEHSVEQVFRNNGEVEFRLEPPGSGYVYVGPVSKLMSKKAFAFGKWKMLEET 252
Db | : | | | | : | | | | | | | | | | | | | | | | | | | | | |
182 GIRSPPGREHSVEYFRNNGGDEVEFRLEPPGSGYVYVGPVSKLMSKKAFAFGKWKLEON 241
OY 253 GGSIIIGTFKAIKRERSSTPKARDPRLPKKPGQTGVSFRKGLRMLDPAISARLSSKLIKLS 312
Db 242 GGSIIIGTFKAIORERKNAKPRDPRLPKKPGQTVGSFRGLRMLDPAISARLSSKLIKLS 301

```

```

OY 313 WKLSITKSEKGYHLTYEPEGVVSIQSRSIYMTVPSYVVAASNIIRLPSVAADALSNEY 372
Db | : | | | | : | | | | | | | | | | | | | | | | | | | | | |
302 WKLSITKLESGYVNLVYPPDGLVYSQSVYVWFVSHVASGLIRLSSAAMNLSKLY 361
OY 373 YPPVAVYIYSYPOAIDERLVDGELKFGQHLFRYQGVETLGTIVSSSLFPPRARGRV 432
Db | : | | | | : | | | | | | | | | | | | | | | | | | | | | |
362 YPPVAVSISYPKVAIRTECLIDGELKFGQLHFRYQGVETLGTIVSSSLFPPRARGRV 421
OY 433 LLNVYIGGAKKPELTSKTEQQIVYVVDRLRKMIRKQADPLVYGVYRVPQALPQFLV 492
Db | : | | | | : | | | | | | | | | | | | | | | | | | | | | |
422 LLNVYIGGSGTWFGLSSEBELVAVDRDLRKMIRKNSHDPDKLVGRRVQAPALPQFLV 481
OY 493 HLDPLSTAKAAMDNDGLEGFLGGNYVAVLGRCEVAGAYEAVSYVGLFSRYAK 548
Db 482 HFDLDTAKSSLSYSSGVYEGFLGGNYVAVYALGRCEVAGAYETAIVANNPBRKYAK 537

```

```

RESULT 7
US-08-808-323-2
: Sequence 2, Application US/08808323
: Patent No. 6018105
: GENERAL INFORMATION:
: APPLICANT: Johnson, Marie
: APPLICANT: Volrath, Sandra
: TITLE OF INVENTION: Promoters from Plant
: TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 601810artis Corporation
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591-9005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/808,323
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/012,705
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/013,612
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/020,003
: FILING DATE: 21-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: CGC 1846
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 537 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-808-323-2

```

Query Match 73.0%; Score 2075; DB 3; Length 537;
 Best Local Similarity 75.4%; Pred. No. 5,7e-198;
 Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

TITLE OF INVENTION: OXIDASE ("PROTOK")
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESSES:
 ADDRESS: No. 6084155artis Corporation
 STREET: 3054 Cornwallis Road
 CITY: Research Triangle Park
 STATE: NC
 COUNTRY: USA
 ZIP: 27709
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/102,420B
 FILING DATE: 22-JUN-1998
 CLASSIFICATION: 800
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 09/059,164
 FILING DATE: 13-APR-1998
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/808,931
 FILING DATE: 28-FEB-1997
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 60/012,705
 FILING DATE: 28-FEB-1996
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 60/013,612
 FILING DATE: 28-FEB-1996
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 60/020,003
 FILING DATE: 21-JUN-1996
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/472,028
 FILING DATE: 06-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
 TELEPHONE: (919) 541-8587
 TELEFAX: (919) 541-8589
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 537 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 US-09-102-420B-2

Query Match 73.0%; Score 2075; DB 3; Length 537;
 Best Local Similarity 75.4%; Pred. No. 5.7e-198;
 Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;
 QY 22 LAFLNRTSFPISPSISKRSNVNCGM---RTRCSYAKDYTPVSSAVDGGPAAEL--DCYI 76
 Db 3 LSLRRTTQSLTSPRSKRV-LRLNYYKRLRRCVYAGPTVGGSSKIRBGGGTTTTPDCVI 61
 QY 77 VAGSISGLCIAOYMSAN---PMLMTEARDRAGNITTVTERDGYLWEEGPNSEFPSPD 132
 Db 62 VEGGISGLICIAALAKTKHPDAAPNLIIVTEAKDRVGNITITREENGFLWEEGPNSEFPSPD 121
 QY 133 MLEMAVDCGLKRDLDVVGPNAPRPVLMKKGLRVPVSKRTDLPFPFDLMSIRGKLRAGGPI 192
 Db 122 MLEMAVDCGLKRDLDVVGPNAPRPVLMKKGLRVPVSKRTDLPFPFDLMSIRGKLRAGGPI 181

QY 193 GLRSPPGHEHSVQFYRRNLIGVEFERLLIEPCSGVYVGDPSKLSMKAAFGVWKLEBT 252
 Db 182 GIRSPPPGERSHVEFYRRNIGDEVEFERLLIEPCSGVYVGDPSKLSMKAAFGVWKLEBQN 241
 QY 253 GGSIIIGTEFKIKERSSTPKAPRDRPRLPKPKGQTQGSFRKGLRMLDPAISARLGSUKLS 312
 Db 242 GGSIIIGTEFKAIQERKKAAPKAPRDRPRLPKPKGQTQGSFRKGLRMLDPAISARLGSUKLS 301
 QY 313 WKLSITRSEKGYHLYETPEPGEVSLQSRSIYMTVPSYASNIILRPLSVAADALSNEY 372
 Db 302 WKLSITRLESGGYNLTYETPDGLVSYQSKVMTVPESHVASSGLRPLSESAANLSTKY 361
 QY 373 YPPGAVLTVSPQAIQDEBLVNGELKRGQLHRPQGVETLGTIYSSLPFNRAPKGRV 432
 Db 362 YPPVAAVLSYPKAIQTECLIDELKGFGLHRPQGVETLGTIYSSLPFNRAAPGRH 421
 QY 433 LLNLYGANKRPELISKTESQVVEVDRLRKMILKPKKODDPLVGVVRWPAITQPLVIG 492
 Db 422 LLNLYGSGTNTGILSKSEGEVAVDRLRKMILKPNSTDEPKLGVYRWPQAITQPLVIG 481
 QY 493 HDLTLSTAKAAMDNGLEGLFLGNGYVGVVALGRVVEGAVYVASEVTGLSRYAYK 548
 Db 482 HFDILDTAKSSLTSSGVEGLFLGNGYVGVVALGRVVEGAVYVETAIYVNNFMSRYAYK 537

RESULT 10
 US-09-071-296-2
 ? Sequence 2, Application US/09071296
 ? Patent No. 6177245
 ? GENERAL INFORMATION:
 ? APPLICANT: Ward, Eric R
 ? TITLE OF INVENTION: Manipulation of Protoporphylinogen
 ? TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms
 ? NUMBER OF SEQUENCES: 12
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Ciba-Geigy Corporation
 ? STREET: 7 Skyline Drive
 ? CITY: Hawthorne
 ? STATE: NY
 ? COUNTRY: USA
 ? ZIP: 10532
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: Patentin Release #1.0, Version #1.25
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/09/071,296
 ? FILING DATE: 06-JUN-1995
 ? CLASSIFICATION:
 ? PRIORITY APPLICATION DATA:
 ? APPLICATION NUMBER: US 08/261,198
 ? FILING DATE: 16-JUN-94
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Elmer, James Scott
 ? REGISTRATION NUMBER: 36,129
 ? REFERENCE/DOCKET NUMBER: CGC 1748/CIP
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: 919-541-8614
 ? TELEFAX: 919-541-8689
 ? INFORMATION FOR SEQ ID NO: 2:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 537 amino acids
 ? TYPE: amino acid
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: Protein
 ? US-09-071-296-2
 Query Match 73.0%; Score 2075; DB 4; Length 537;
 Best Local Similarity 75.4%; Pred. No. 5.7e-198;

```
Matches 404; Mismatches 73; Gaps 4;
QY 22 LAFLNRSPFSSSSIKSNVNCNM--RTRCSVADKTVYSSAVDGGPAEL--DCVY 76
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 3 LSLARPTQSLPLPSSKRN-LRLWYKRLRLRCVAGGPPVSSKIGGGTITTDVY 61
QY 77 VGAGTSGICLQAVMSAN--PILAVTEARDRAGNITTVERDDGYMEGPPNSFQPSDF 132
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 62 VGGTSGICLQALATKRPDAAPNLYEAKDRVGMNITRREBGEFLMEBGGPNSFQPSDF 121
QY 133 MLEMAVDCGLADDLVLDGPNAPRFLWMLKRLRVPVSKLTLDLPPFDLMSIGKLRAAGPPT 192
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 122 MLITMVDGSLDDLVLDGPTAPRPRVYLNWVKRLRYPVSKLTDLPPFDLMSIGKLRAAGPPT 181
QY 193 GIRPSPGHESSVEQVFNKLGVEFERLLEPFCSSGVYVDDPSKLMSKAAFGKVMKLEET 252
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 182 GIRPSPGHESSVEEYFVRNLDGEFERLLEPFCSSGVYVADGPKSLMSKAAFGKVMKLEEQN 241
QY 253 GSSITGGTFKAIKRSPRPAKRPDRPLPKPKGGVGSFRKGLRMLPAISARLGSKLTKS 312
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 242 GSSITGGTFKAIQKRNPKAERDPRLPKPQGTGVSFRKGLRMLPAISARLGSKLVKLS 301
QY 313 WLSSTPTSEKGVHLYETPEGVVSLQSRVIVTVPVYVANLRLPPLVAADALSNEFY 372
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 302 WLSGITKLESGGYNLVYETPDGLVYVSKVVMVTPSHVASSGLRPLSSANALSILY 361
QY 373 YRPGAAYVTSYRQAIKRLRERLVDDELKFGQLHRRTGVEETLGTIYSSSLEFPARAKGRY 432
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 362 YRPAVAASVSTPKAIRECLIDELKFGQLHRRTGVEETLGTIYSSSLEFPARAPGRY 421
QY 433 LLNLTGKAKNPELISKTESQLVEVDRDLRKLMLIKKADPVLVGVGRVQAIPOFLVG 492
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 422 LLNLTGSTMNTGILSSSEGLVAVDRDLRKLMLIKRNSDPLDKLGRVWQAIPOFLVG 481
QY 493 HLDLITAKAMNDNGLEGLFGLGNVYSGVALRCVGAEVASEVTGFLSRVAYK 548
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 482 HFDILDITAKSSLSLSSGEGYGLFGLGNVYSGVALRCVGAETAIEVNMMFSRVAAYK 537
RESULT 11
US-08-808-931-12
? Sequence 12, Application US/08808931
? Patent No. 5939602
? GENERAL INFORMATION:
? APPLICANT: Johnson, Sandra
? APPLICANT: Johnson, Marie
? APPLICANT: Potter, Sharon
? APPLICANT: Ward, Eric
? APPLICANT: Helfetz, Peter
? TITLE OF INVENTION: DNA Molecules Encoding Plant
? TITLE OF INVENTION: Protoptorphyrinogen Oxidase and Inhibitor-Resistant Mutants
? NUMBER OF SEQUENCES: 35
? CORRESPONDENCE ADDRESS:
? ADDRESS: No. 5939602artis Corporation
? STREET: 520 White Plains Road, P.O. Box 2005
? CITY: Tarrytown
? STATE: NY
? COUNTRY: USA
? ZIP: 10591-9005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/808,931
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/012,705
? FILING DATE: 28-FEB-1996
? PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 543 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-931-12
Query Match 72.3%; Score 2055; DB 2; Length 543;
Best Local Similarity 78.1%; Pred. No. 5,7e-196;
Matches 392; Conservative 44; Mismatches 60; Indels 6; Gaps 3;
QY 50 RGSVAKYTV-PSSAVDGGPRAEDLYVNGAGTSGICLQAVMSAN--PILAVTEARDA 106
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 45 RCSIARSBSTASPKRTRSA--VDCVAVVGGVSGICLQALATKRNANAVVTEARDRY 101
QY 107 GGNITVVERDGYLMEEGPNSFQPSDPMLEMAVDCGLKDDLVLGDPNAPRFLVLMKRLRVP 166
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 102 GGNITVVERDGYLMEEGPNSFQPSDPMLEMAVDCGLKDDLVLGDPNAPRFLVLMKRLRVP 161
QY 167 PSKLTDPFPLDMSITPKLRAAGSPVGLRSPGHESSVEQVFNKLGVEFERLLEPFC 226
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 162 PGKLTDPFPLDMSITPKLRAAGSPVGLRSPGHESSVEQVFNKLGVEFERLLEPFC 221
QY 227 SGVYVDDPSKLSMKAAFGKVMKLEETGSIIGTFFAKIKRSPRPAKRPDRPLPKPQGT 286
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 222 SGVYVDDPSKLSMKAAFGKVMKLEETGSIIGTFFAKIKRSPRPAKRPDRPLPKPQGT 281
QY 287 VGSFRKGLRMLPAISARLGSKLTKSITKSEKGVHLYETPEGVVSLQSRVIVM 346
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 282 VGSFRKGLRMLPAISARLGSKLTKSITKSEKGVHLYETPEGVVSLQSRVIVM 341
QY 347 TFPSTYVANLRLPPLVAADALSNEFYYPVAVTISYRQAIKRLRERLVDGELKFGQLHP 406
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 342 TFPSTYVANLRLPPLVAADALSNEFYYPVAVTISYRQAIKRLRERLVDGELKFGQLHP 401
QY 407 RFGVEFLGTIYSSSLEFPARAPKGRVLLNLTGAKNPELISKTESQLVEVDRDLRKL 466
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 402 RSQGVETLGTIYSSSLEFPARAPKGRVLLNLTGAKNPELISKTESQLVEVDRDLRKL 461
QY 467 IKPRAADPLVVGVWVMPQAIPOFLVHLDLITAKAMNDNGLEGLFGLGNVYSGVALGR 526
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 462 INPAAQDPFVGVGRVWVMPQAIPOFLVHLDLITAKAMNDNGLEGLFGLGNVYSGVALGR 521
QY 527 CVEGAYEVAEVTGFLSRVAYK 548
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 522 CVEGAYEVAEVTGFLSRVAYK 543
RESULT 12
US-08-808-923-12
? Sequence 12, Application US/08808923
? Patent No. 6018105
? GENERAL INFORMATION:
? APPLICANT: Johnson, Marie
? APPLICANT: Voltzath, Sandra
? APPLICANT: Ward, Eric
? TITLE OF INVENTION: Promoters from Plant
? TITLE OF INVENTION: Protoptorphyrinogen Oxidase Genes
? NUMBER OF SEQUENCES: 26
```



```

? Patent No. 5939602
? GENERAL INFORMATION:
? APPLICANT: Voltrath, Sandra
? APPLICANT: Johnson, Marie
? APPLICANT: Potter, Sharon
? APPLICANT: Ward, Eric
? APPLICANT: Heiletz, Peter
? TITLE OF INVENTION: DNA Molecules Encoding Plant
? TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
? NUMBER OF SEQUENCES: 35
? CORRESPONDENCE ADDRESS:
? ADDRESSER: No. 5939602artlis Corporation
? STREET: 520 White Plains Road, P.O. Box 2005
? CITY: Tarrytown
? STATE: NY
? COUNTRY: USA
? ZIP: 10591-9005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/808,931
? FILING DATE:
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/012,705
? FILING DATE: 28-FEB-1996
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/013,612
? FILING DATE: 28-FEB-1996
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 60/020,003
? FILING DATE: 21-JUN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Meigs, J. Timothy
? REGISTRATION NUMBER: 38,241
? REFERENCE/DOCKET NUMBER: CGC 1847
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (919) 541-8587
? TELEFAX: (919) 541-8689
? INFORMATION FOR SEQ ID NO: 20:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 536 amino acids
? TYPE: amino acid
? STRANDEDNESS: not relevant
? TOPOLOGY: not relevant
? MOLECULE TYPE: protein
? US-08-808-931-20

```

```

? Query Match 70.6%; Score 2006.5; DB 2; Length 536;
? Best Local Similarity 73.1%; Pred. No. 3.8e-191;
? Matches 391; Conservative 52; Mismatches 83; Indels 9; Gaps 4;

```

```

QY 22 IAFINRTFSI-PESSISKRNNGWRTKRSYAKDVTVPSSAVDDGPAE--LDCVTV 77
DB 3 LSLIRQPLSPFSPNPPRR-SRPYKPLMLKCSVSGSVSSSTIEGGGGGKTVTADCVTV 61
QY 78 GAGISGLCIAQVMSANYP---NLMVTEARDRAGNITTVVERDGYLWEEGPNPSPSDPM 133
DB 62 GGGISGLCIAQALVTKRHDAAKNVYVTEAKDRVGNITTRREGDGLWEEGPNPSPSDPM 121
QY 134 LFMAYVDCGKDKDLVGDNDAPPVLMKGLRPPVPSKLLDLPPFDLMSIPGKLRAGGPTG 193
DB 122 LFMVVDSSGLAKDLDVGDPTAPRPFVIMNGKLRPVPSKLLDLPPFDLMSIGKIRAGGALG 181
QY 194 LRSPPGHEESVEQFVRRNLGGEVFERLIEPFCGYYVGDPSKLSMKAAPGKVMKLEETG 253
DB 182 IRSPPGRESEVEEFRNRNLGDEVEFERLIEPFCGYYVADPAPAKLSMKAAPGKVMKLEENG 241

```

```

QY 254 GSIIIGTFPKAIKERSSTPKAPRDPRLPKKGGTVGSPFKKGLRMLPDAISARLGSKTLKLSW 313
DB 242 GSIIIGAFKAIQAKNKAPRTTRDPLPKKGGTVGSPFKKGLTMLPEAISARLGDKVVSM 301
QY 314 KLSSTIKSEKGGHLLTYEPPEGVSLGSRSLVMYVPSVVAANIIRPISVAADALSNFY 373
DB 302 KLSSTIKLASGESLTYEPPEGVSLGSRSLVMYVPSVVAASLLRPLSDSAAERLSKLY 361
QY 374 PPGAANTISYPOFAIRDERLVDGELKGFGLHPTGVEITLGTIYSSSLPNNRPAKGRVL 433
DB 362 PVAAVSISYAKAIRSECLIDGELKGFGLHPTGVEITLGTIYSSSLPNNRPAKGRVL 421
QY 434 LMYTIGAKKPELSTESQVLEVYVDRDLRKLKPKAQDDLVVGVYVWQALPQFLVGH 493
DB 422 LLMVIGGANTVGLSKSEBELVEAVDRDLRKLKPKSTDPVLVGVKLMQALPQFLIGH 481
QY 494 LDFLSTAKAAMNDNGLEGLFLGNTVVSVALGRVCEGAYEVAASEVTFGLSRVAVK 548
DB 482 IDLVDAKAKASLSSSGHGLFLGNTVYVAGVALGRVCEGAYEVAASEVTFGLSRVAVK 536

```

```

Search completed: July 3, 2001, 10:31:54
Job time: 1913 sec

```



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:33:07 ; Search time 34.09 Seconds
(without alignments)
550,660 Million cell updates/sec

Title: US-09-508-418-2

Perfect score: 2844
Sequence: 1 MTRTPPIANNPNTFRHOSSSS.....EGAEVASEVTGFLSRVAVK 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	2821	99.2	548	1 PPOC_TOBAC	O24163 nicotiana t
2	2075	73.0	537	1 PPOC_ARATH	P55826 arabidopsis
3	541	19.0	471	1 PPOC_MYXXA	P56601 myxococcus
4	505	17.8	504	1 PPOM_TOBAC	O24164 nicotiana t
5	485	17.1	470	1 PPOX_BACSU	P32397 bacillus su
6	373.5	13.2	527	1 PPOX_PROFU	O32634 protoplombac
7	363.5	12.8	477	1 PPOX_HUMAN	P50336 homo sapien
8	362.5	12.7	477	1 PPOX_MOUSE	P51175 mus muscucu
9	302.5	10.6	490	1 YAM7_SCHPO	Q10062 schizosacch
10	221	7.8	426	1 PPOX_MYCTU	O53230 mycobacteri
11	216.5	7.6	451	1 PPOX_MYCLE	O50008 mycobacteri
12	189	6.6	539	1 PPOX_YEAST	P40012 saccharomyc
13	141	5.0	448	1 AOPH_MYCTU	O53320 mycobacteri
14	139	4.9	570	1 CRTI_MARPS	Q40406 narcissus s
15	134	4.7	522	1 AOPF_ONCAM	P49353 oncorhynchu
16	132.5	4.7	527	1 AOPF_BOVIN	P21398 bos taurus
17	131	4.6	566	1 CRTI_ARATH	O07356 arabidopsis
18	125.5	4.4	417	1 YVAB_RH1N3	P55349 rhizobium s
19	125.5	4.4	472	1 CRTI_SYNV3	P29273 synchocyst
20	125.5	4.4	489	1 CBPI_GANAL	P31325 candida alb
21	123.5	4.3	526	1 AOPF_ARAT	F21596 rattus norv
22	123	4.3	582	1 CRTI_CAPAN	P80093 capsicum an
23	122	4.3	471	1 V782_SYNV3	Q55629 synchocyst
24	119.5	4.2	478	1 PPO_MICRU	P40974 micrococcus
25	118	4.1	571	1 CRTI_MAIZE	P49086 zea mays (m
26	117.5	4.1	469	1 P49_STRLI	P06108 streptomyce
27	116	4.1	583	1 CRTI_LYCES	P28554 lycopersico
28	114.5	4.0	1106	1 TNA7_RAT	O63258 rattus norv
29	114	4.0	531	1 ACHC_ACHHU	P33903 achactina fu
30	112.5	4.0	570	1 CRTI_SOYBN	P28553 glycine max
31	111	3.9	508	1 CRTI_STRSE	P54971 streptomyce
32	109.5	3.9	527	1 AOPF_HUMAN	P21397 homo sapien
33	109.5	3.9	908	1 SRCA_RABIT	P13666 oryctolagus

RESULT ID	Score	Query Match	Query Length	DB ID	Description
34	109	3.8	1181	1 TNA7_HUMAN	Q13663 homo sapien
35	108.5	3.8	1062	1 CERU_MOUSE	O61147 mus muscucu
36	108	3.8	504	1 A37C_DROME	P18487 drosophila
37	107.5	3.8	516	1 ER13_ARATH	O65403 arabidopsis
38	107	3.8	405	1 SNOX_CORSI	P40875 corynebacte
39	107	3.8	1179	1 TNA7_MOUSE	O61738 mus muscucu
40	106.5	3.7	630	1 FGI1_MOUSE	O09046 mus muscucu
41	105.5	3.7	872	1 IP3L_RAT	P42335 rattus norv
42	105	3.7	755	1 TR2M_AGRRT3	P25017 agrobacteri
43	104.5	3.7	466	1 AMD_PSEPU	O69768 pseudomonas
44	104.5	3.7	517	1 ER12_ARATH	O65402 arabidopsis
45	104.5	3.7	524	1 CRTI_RHOCA	P17054 rhodobacter

ALIGNMENTS

RESULT 1
ID PPOC_TOBAC
AC 024163; STANDARD; PRT; 548 AA.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO I)
DE (PROTOPORPHYRINOGEN IX OXIDASE ISOZYME I) (PPX I).
GN PPXI.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV_SRI;
RX MEDLINE=97385200; PubMed=9238074;
RA Lemnitova I., Kruse E., Mock H.-P., Grimm B.;
RT Cloning and characterization of a plastidial and a mitochondrial
RT isoform of tobacco protoporphyrinogen IX oxidase. "
RL Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997).
CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC IX TO FORM PROTOPORPHYRIN IX.
CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC IX + H(2)O(2).
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: PENTULATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS, AND
CC CHLOROPHYLL SYNTHESIS.
CC -1- SUBUNIT: HOMODIMER; CONTAINS ONE FAD PER HOMODIMER (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREMATURE LEAVES.
CC DECREASED EXPRESSION IN OLDEST LEAVES. EXPRESSED AT VERY LOW LEVEL
CC IN ROOTS.
CC -1- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL
CC EXPRESSION IN THE DARK PERIOD.
CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: Y13465; CAA73865.1; -
KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW Transit peptide; Flavoprotein; FAD; Chloroplast;
KW Chlorophyll biosynthesis.
FT TRANSLIT 1
FT CHAIN 51 548 CHLOROPLAST (POTENTIAL).
FT NP_BIND 78 83 FAD (POTENTIAL).
FT SEQUENCE 548 AA; 59230 MW; 668992E78FB8A3EE30 CRC64;

Query Match 99.3% Score 2821: DB 1: Length 548;
 Best Local Similarity 99.3% Pred. No. 1,4e-205;
 Matches 544; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY	1	MTPPIANHPNIFTHQSSSSPLAFLNRTSFTPFSSISRNSVNCNGMRTKCSVAKDYTP	60
DB	1	MTPPIANHPNIFTHQSSSSPLAFLNRTSFTPFSSISRNSVNCNGMRTKCSVAKDYTP	60
OY	61	SSAVNGGPAALDQVYAGTSSGICIAOVMSANYPNIMVMEARBRACGNITTPEROCYIM	120
DB	61	SSAVNGGPAALDQVYAGTSSGICIAOVMSANYPNIMVMEARBRACGNITTPEROCYIM	120
OY	121	EEGPNFQSDPMLTMAVDCGLKDDVYLGDENAPRFLVMKPKLPPVPSKLTLDLPPFDLMS	180
DB	121	EEGPNFQSDPMLTMAVDCGLKDDVYLGDENAPRFLVMKPKLPPVPSKLTLDLPPFDLMS	180
OY	181	IPGKIRAGFPIGLRPSRPGHESEVQFVRNLDGVEPERLIEPFGCGVYVGPSSKLSMK	240
DB	181	IPGKIRAGFPIGLRPSRPGHESEVQFVRNLDGVEPERLIEPFGCGVYVGPSSKLSMK	240
OY	241	AAPGKWKLEETGGSIIGETPKAIKERSSTPKAPDRDRLPKPKGQVGSFRKGLRMLPDA	300
DB	241	AAPGKWKLEETGGSIIGETPKAIKERSSTPKAPDRDRLPKPKGQVGSFRKGLRMLPDA	300
OY	301	ISARLGSKLKLSKMSLSTTKSEKGYHLTTEPPGVYSLQSRSLVMFTVPSYASNLIRPL	360
DB	301	ISARLGSKLKLSKMSLSTTKSEKGYHLTTEPPGVYSLQSRSLVMFTVPSYASNLIRPL	360
OY	361	SVAAADALSNFYPPVGAVTISYQEAIRDEIRLVDGELKFGQGLHPHTQGVETLGIYSS	420
DB	361	SVAAADALSNFYPPVGAVTISYQEAIRDEIRLVDGELKFGQGLHPHTQGVETLGIYSS	420
OY	421	SLEPFRARAKGVLLIINTIGCAKKNPEIISKTESQVLEVVYDRLKMLIKRKADPLVYGR	480
DB	421	SLEPFRARAKGVLLIINTIGCAKKNPEIISKTESQVLEVVYDRLKMLIKRKADPLVYGR	480
OY	481	WVPOAIPQFLVGHLDLSTAKAANNNGLELFLGNYVSGVALRCEVAGYEVASEVYNG	540
DB	481	WVPOAIPQFLVGHLDLSTAKAANNNGLELFLGNYVSGVALRCEVAGYEVASEVYNG	540
OY	541	FLSRVAYK 548	
DB	541	FLSRVAYK 548	

RESULT 2
 PROC_ARATH STANDARD; PRT: 537 AA.

ID	PROC_ARATH	STANDARD	PRT	537 AA.
AC	P59826;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	PROTOPORPHRYRINOEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO).			
GN	POX OR AT4G01690 OR T15816.13			
OS	Arabidopsis thaliana (Mouse-ear cress).			
CC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
CC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosoids II;			
CC	Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBITaxID=3702;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97136707; Pubmed=8982084;			
RA	Marita S., Tanaka R., Ito T., Okada K., Taketani S., Inokuchi H.;			
RT	"Molecular cloning and characterization of a cDNA that encodes			
RN	protoporphyrinogen oxidase of Arabidopsis thaliana.";			
RL	Gene 182:169-175(1996).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA.			
RX	MEDLINE=20083488; Pubmed=10617198;			
RA	Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.;			

RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Anstorge W., Brandt P., Griwell L., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaler B., Mache R., Muller M.,
 RA Kreis M., Delseny M., Puljdomenich P., Watson M., Schmidheini T.,
 RA Reichelt B., Portelelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbusche F.,
 RA Breken M., Weljens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Welfzenger T., Bothe G., Rempesberger U., Hilbert H., Braun M.,
 RA Holzner E., Brandt A., Peters S., van Staveren M., Dirse W.,
 RA Woolman P., Klein Lambhorst R., Kose M., Hauf J., Koeltter P.,
 RA Beneser S., Hempel S., Feldpausch M., Lambers S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quall M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McRay K., Mayes R.,
 RA Pettett A., Raizendream M.-A., Lyne M., Bees V., Rechmann S.,
 RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderrath K., Dauner D., Herzl A.,
 RA Neumann S., Argiridou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedori F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Bagnues M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bleik C.,
 RA Heijnen D., Haase D., Lemcke K., Mews H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Basilde M., Habermann K.,
 RA Parnell L., Dediha N., Guoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Thiraidh J.,
 RA Stonerkeng T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Larrelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mink P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Krimmer J., Fulton L., Marzils E., Dante M., Pegin K., Hillier L.,
 RA Nelson J., Spiehl J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
 RA Antonolou B., Zidanic M., Strong C., Sun H., Lamer B., Jordan C.,
 RA Ma P., Zhong Y., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swebby I.K., O'Shaunnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lochi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., Mccombe W.R.;"

RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:769-777(1999).

CC -I- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHRYRINOEN
 IX TO FORM PROTOPORPHRYRIN IX.
 CC -I- CATALYTIC ACTIVITY: PROTOPORPHRYRINOEN-IX + O(2) - PROTOPORPHRYRIN-
 IX + H(2)O(2). (BY SIMILARITY).
 CC -I- COFACTOR: FAD (BY SIMILARITY).
 CC -I- PATHWAY: PENTULIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -I- SUBCELLULAR LOCATION: CHLOROPLAST (POTENTIAL).
 CC -I- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE LEAVES AND AT
 LOW LEVELS IN THE ROOTS AND FLORAL BUDS.
 CC -I- SIMILARITY: BELONGS TO THE PROTOPORPHRYRINOEN OXIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D83139; BA11820.1;
 DR EMBL: AL161492; CAB7739.1;
 KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
 KW Flavoprotein; FAD; Chloroplast; Transil peptide;
 FT TRANSLIT 1 ?
 FT CHAIN 2 ?
 FT NP_BIND 63 537
 FT SEQUENCE 537 AA; 57695 MW; C125B0E5D13187D5 CRC64;

Query Match 73.0%; Score 2075; DB 1; Length 537;
 Best Local Similarity 75.4%; Pred. No. 3,1e-149;
 Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

22 LAFILRTSIPSSSTSKNSVNCNM---PRRCVAKRYIYPPSSAVDGGPAEL--DCVI 76
 3 LSLLRPTQSLIPSPSKRN-LKLNRYKPLRLRCVAGGPTVSSSKIEGGGTTTDDVI 61

77 VGAGISGLCIAQVMSANY-----PMLMTEARDRAGNITTVTERDGYLWEEGPNPSPDP 132
 62 VGGISGLICIAQALATKHPDAAPNLIYEAADRNGNITTEBNGFTLWEEPNPSPDP 121

133 MLTAVDGLKDDLVLGPNAAPRYLVKGGKLRPVPSKLTLDLPPFDLMSIPKLRAGFPI 192
 122 MLTAVDGLKDDLVLGPNAAPRYLVKGGKLRPVPSKLTLDLPPFDLMSIPKLRAGFPI 181

193 GLRPSPEHSESVBOFVRRNLTGGEVEERLTPEFCSGVYVGGPDKSLMKAAGKWKLEET 252
 182 GIRESPPEBRESVEFRRNLTGDEVEFERLTPEFCSGVYVGGPDKSLMKAAGKWKLEON 241

253 GGSIIIGTFKAIKERSSTPKAPRDLRPPKQGTGVSFRKGLRMLPDAISARLSKILKS 312
 242 GGSIIIGTFKAIKERSKNAKAPRDLRPPKQGTGVSFRKGLRMLPDAISARLSKILKS 301

313 WKSITSEKGGYHLTYEPREGVYSLQSNITVTFVPSYVANSILRPLSVAADALSIFY 372
 302 WKSITSEKGGYHLTYEPREGVYSLQSNITVTFVPSYVANSILRPLSVAADALSIFY 361

373 YPPVAVTISYPOAIRDERLVDEGLKGFQGLHPRTOGVEITGIVYSSSLFPNAPKRGV 432
 362 YPPVAVTISYPOAIRDERLVDEGLKGFQGLHPRTOGVEITGIVYSSSLFPNAPKRGV 421

433 LLNVTGKAKNPELISKTESQVLEVVDRDLRKMILKPKADPVLVYGVVWVPOAIPQFLV 492
 422 LLNVTGKAKNPELISKTESQVLEVVDRDLRKMILKPKADPVLVYGVVWVPOAIPQFLV 481

493 HLDITSTAKAMNDNGLEGLFELGGNNVYSGVALGRCVGAAYEVAEYVGFISRAYK 548
 482 HFDLIDTRAKSSITSSYEGLEFLGNNVYSGVALGRCVGAAYEVAEYVGFISRAYK 537

RESULT 3
 PROX_MYXXX STANDARD; PRT; 471 AA.

AC P56601: 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
 GN HEMK.
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 OC Myxococcales; Cystobacteriales; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92192453; PubMed=1372277;
 RA Li S., Lee B.U., Shimkets L.J.;
 RT "csga expression entrains Myxococcus xanthus development.";
 RN Genes Dev. 6:401-410(1992).
 RL [2]
 RP REVISION: AND CHARACTERIZATION.
 RP MEDLINE=96224079; PubMed=8621504;
 RA Dailley H.A., Dailley T.A.;
 RT "Protoporphyrinogen oxidase of Myxococcus xanthus. Expression,
 RT purification, and characterization of the cloned enzyme.";
 RL J. Biol. Chem. 271:8714-8718(1996).
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 CC IX TO FORM PROTOPORPHYRIN IX. MAY HAVE AN ESSENTIAL EXTRACELLULAR
 CC FUNCTION DURING DEVELOPMENT POSSIBLY AS A PHEROMONE.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2).

-1- COFACTOR: CONTAINS ONE FAD PER HOMO-DIMER.
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND Porphyrin BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMO-DIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT AT POSITION 127.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; W73709; NOT_ANNOTATED; CDS.
 DR EMBL; AF098938; AAD13609.1; -
 KW Porphyrin biosynthesis; Oxidoreductase; Flavoprotein; FAD;
 KW Heme biosynthesis.
 FT NE_BIND 16
 FT SEQUENCE 471 AA; 49387 MW; 2E9909D464F81515 CRC64;

Query Match 19.0%; Score 541; DB 1; Length 471;
 Best Local Similarity 31.7%; Pred. No. 1.9e-33;
 Matches 150; Conservative 84; Mismatches 207; Indels 32; Gaps 12;

72 LDCVYVAGISGLCIAQVMSANYPNLMTTEARDRAGNITTVTERDGYLWEEGPNPSPDP 131
 10 MNVAVVGGISGLIAVANHILRSRGTDAVLLLESSARLGGAVGTHALGYLVEQGPNPSPDR 69

132 PML-TMAVDCGLKDDLVLDGPNAP-REVLMTKGLRVPKLTLDLPPFDLMSIPKLRAGF 189
 70 PARALAAALNLEGRRAADPAKRRVYTRGKLRVSPSPAFPLASDILPLGARLRVA- 128

190 GPTGLRPSPEHSESVBOFVRRNLTGGEVEERLTPEFCSGVYVGGPDKSLMKAAGKWKLE 249
 129 GELFSRRAEAGEVDSLAAGRRLHGRATQVLLDAVQGTIVYAGVDEQSLVAATPFMIVKM 188

250 EETGGSIIGTFKAIKERSSTPKAPRDLRPPKQGTGVSFRKGLRMLPDAISARLSKILKS 305
 189 EREHNSLILGALNA-----QKARQALPAGTAPKLSGALSTFDGGIQLVLDLMAASL 241

306 GSKLTKLMSITSEKGGYHLTYEPREGVYSLQSNITVTFVPSYVANSILRPLSVAADALSIFY 365
 242 GDAAHVAGARVEGLAR-EDSGWRLIIEHGRRAELSAQVYVLAAPAAHAKLRPLDLDLA 300

366 DALSNFTYPPVGAVTISYPOAIRDERLVDEGL---KGFQGLHPRTOGVEITGIVYSSSL 422
 301 ALVAGIAYVAPRIVVHLGFPA-----GTLPPADGDFLVPAAEQRRMGLIHAISTT 350

423 FPNRPAKGVLLNVTGKAKNPELISKTESQVLEVVDRDLRKMILKPKADPVLVYGVVWV 482
 351 FPNRPAKGVLLNVTGKAKNPELISKTESQVLEVVDRDLRKMILKPKADPVLVYGVVWV 408

483 PQAIDQFLVGHLDITLSTAKAMNDNGLEGLFELGGNNVYSGVALGRCVGAAYEVA 535
 409 PGIPIQYNLGHLEVAALDAAL--QRLEGLHILGNAYGVGINDCIRNAQAOLA 459

RESULT 4
 PROM_TOBAC STANDARD; PRT; 504 AA.

AC O24164:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE, MITOCHONDRIAL (EC 1.3.3.4) (PPO II)
 DE (PROTOPORPHYRINOGEN IX OXIDASE ISOZYME II) (Ppx II) (Px-2).
 GN PPK11 OR PPO2.
 OS Nicotiana glauca (Common tobacco).
 OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

```

OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanales; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SRI.
RX MEDLINE=97385200; PubMed=9238074;
RA Lermontova I., Kruse E., Mock H.-P., Grimm B.;
RT Cloning and characterization of a plastidal and a mitochondrial
RT isoform of tobacco protoporphyrinogen IX oxidase.
RL Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. Samsun NN.
RX PubMed=9808719;
RA Watanabe N., Che F., Iwano M., Takayama S., Nakano T., Yoshida S.,
RA Tsogai A.;
RT Molecular characterization of photomixotrophic cultured tobacco cells
RT resistant to protoporphyrinogen oxidase-inhibiting herbicides.
RL Plant Physiol. 118:751-758(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SRI.
RA Horikoshi M., Mametsuka K., Hirooka T.;
RT The molecular basis of photobleaching herbicide resistance in
RT tobacco.
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
RC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
RC IX TO FORM PROTOPORPHYRIN IX.
CC -1- FUNCTION: PROVIDES PRECURSOR FOR THE MITOCHONDRIAL AND PLASTIDIC
CC HEME SYNTHESIS AND THE PREDOMINANT CHLOROPHYLL SYNTHESIS IN
CC PLASTIDS.
CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) =
CC PROTOPORPHYRIN-IX + H(2)O(2).
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- ENZYME REGULATION: INHIBITED BY THE HERBICIDE ACIFLUORFEN.
CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRION.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREMATURE LEAVES.
CC DECREASED EXPRESSION IN OLDEST LEAVES. ALSO DETECTED IN ROOTS.
CC -1- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL
CC EXPRESSION IN THE DARK PERIOD.
CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC
DR EMBL; Y13466; CAAT7866.1; ..
DR EMBL; AB020500; BAA34712.1; ..
DR EMBL; AF044129; AAD0291.1; ..
DR Interpro: IPR002937; ..
DR Pfam: pf01593; Amino oxidase; 1.
KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW Flavoprotein; FAD; Mitochondrion.
FT NP_BIND 20 25 FAD (POTENTIAL).
FT SEQUENCE 504 AA; 55407 MW; B85B55EC81DC00A CRC64;

```

```

Query Match 17.88; Score 505; DB 1; Length 504;
Best Local Similarity 28.88; Pred. No. 1.1e-10;
Matches 149; Conservative 100; Mismatches 208; Indels 60; Gaps 17;
OY 60 PSSAVDGGPAALDQVIYAGISGICIAOVMSANYPNIMLWTEARDRAGNTITVERDGYL 119
  11: 1:1: :||||| 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
DB 3 PSAGEDKHSMAK-HVAVYIGAGVSGLAAYAKKIKHGLVWTFEABGCRKGRSVDGQGLD 61
  1:1:1:1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
OY 120 WEEGNSFQSDPMLTGMADV-CGLKDLVLGDDPNAKPVLMKGLRVPVSKLTDLPDFDL 178
  1:1:1:1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:

```

```

Db 62 WDEGANMTTESEGDVTFILDLSGLREKQKPELSONRKRIRYRANGNVPDLLPSPIDLIKSNP 121
OY 179 MSIRPKLRAGFGPI-----GLRSPFPGHESVEQVFYRNRLGGEFFELTEPFCOSGYYVD 233
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
Db 122 LSTGSKLQMLLEPTLMKRNKRLSQQVSDSH-ESVSGEFGFRRHFGKEVVDLIDFVAGTCGGD 180
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
OY 234 PSKLSMKAAPFGKVMKLEPTGGSIIGTEFKA-IKERSSTPKAPRPRLPKPGQTVGSFRR 292
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
Db 181 PSDLSMHHSPFELMNLKRRFSGVILGAIKRLSPKRNKKGCPKPTSNKRNKRGSGFSLG 240
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
OY 293 GLRMPLDPAI-----SARLGSK-IKL-----SMKLSITSEKGGCHLYETP 333
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
Db 241 GNGTLTDAICKDRLREDELRLNSRVLELSCSETEDSALDMSIISASPHKQSESESD- 298
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
OY 334 EGVVSLQRSIVMPPVSVVANNIIRPLSLVAADAAMSNFY-----YPPGAVTISYPOAIR 389
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
Db 299 -----AVIMTAPLDCVKS--KIARRGNPLNITPEVDVYPLSVVITTTKRNVK 347
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
OY 390 DERLVDGELKGFOLHP---RTQGVETLGTIYSSSLPFPNRRAPKRVLLNIYGGAKNPEI 446
  1:1:1:1: 1:1:1:1: 1:1:1:1: 1:1:1:1: 1:1:1:1: 1:1:1:1: 1:1:1:1: 1:1:1:1: 1:1:1:1: 1:1:1:1:
Db 348 -----YPLEEGFYLVPSKEQHGLTKLTGLTFSSMPPDRAPNNVVLYLTFVGGSRREL 401
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
OY 447 LSKTESQLVEYVDRDLRKMILKPKAQDPLVYGVWVPAIPLQFLVGH-LDTLSTAKAAM 505
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
Db 402 AKASRTEIKEIYTPDKQLL--GAEGEPEYVNNHLVWSKAEP--LYGHNYDSVLDALDKME 457
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
OY 506 DNGLEGFLGNGVYSVALGRCVBAGYEVASEVYGF 542
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
Db 458 KN-DPGLFYLAGNHKRGSLVGNKALSSGCAADLVISTYL 493
  1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
RESULTS
PROX_BACSU STANDARD: PRT: 470 AA.
ID P32397; STANDBY
AC P32397; 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DR 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PRO).
GN HEWY OR HEMG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=158;
RC Nodack M.A., Terpestra P., Holsappel S., Venema G., Bron S.;
RA Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93094140; PubMed=1459957;
RA Hansson M., Hederstedt L.;
RT Cloning and characterization of the Bacillus subtilis hemEh gene
RT cluster, which encodes prothème IX biosynthetic enzymes.
RL J. Bacteriol. 174:8081-8093(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=158;
RC Nodack M.A., Terpestra P., Holsappel S., Venema G., Bron S.;
RA Dalley T.A., Meissner P., Dalley H.A.;
RT *Expression of a cloned protoporphyrinogen oxidase.
RL J. Biol. Chem. 269:813-815(1994).
RN [1]
RP FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
RN IX TO FORM PROTOPORPHYRIN IX. ALSO OXIDIZES THE PATHWAY
RN INTERMEDIATE COPROPORPHYRINOGEN III.
CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC IX + H(2)O(2).
CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC -1- PATHWAY: PENULTIMATE STEP IN PORPHYRIN BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M97208; AAA22519.1; -
 DR EMBL; Y14083; CAAT74520.1; -
 DR EMBL; Z99109; CAB12854.1; -
 DR PIR; D47045; D47045.
 DR HSP; F54725; 10BA.
 DR Subtilist; BG10431; hemX.
 KW Porphylin biosynthesis; Oxidoreductase; Flavoprotein; FAD.
 FT NP_BIND 12 17 FAD (POTENTIAL).
 SO SEQUENCE 470 AA; 51203 MW; 95CC4E847686D2E CRC64;

Query Match 17.1%; Score 485; DB 1; Length 470;
 Best Local Similarity 28.7%; Pred. No. 3.2e-29;
 Matches 139; Conservative 98; Mismatches 206; Indels 42; Gaps 15;

OY 75 YVYGAGISGLCIA-----QVMSANP-NLMTVEARDRAGNITTVYERDGYLMEEGPNSF- 127
 Db 9 YIIIGGITGIAAAYFKWEKIKKKNPLLETLTVAESPFGKIQYKCKDYIIRGPDSEFL 68
 OY 128 --QPSPMLTMAVDCGLKBDLVLDGPNAPRVIY-----KGLRIPSKLTLPDFPD 177
 Db 69 ERKKSAPQLVK--DLGL-ELLVNNATGQSYLVNRTLHPMKGAVMGITPTLAFPVSTG 125
 OY 178 LMSIPEKLRAGFGPIGLRPSRPGHEESYQVFRNMLGSEVFBRLLTEPCSGVYVGDPSKL 237
 Db 126 LESLSGKARAAMDPI-LPASKTRKDCQSLGEPFRFRVGDVEVNLLEPLLSGIYAGDIDKL 184
 OY 238 SMKAAGKXWKLLEFNGGSIIGGTFKAIKRSSSTPKARPRPRPKRGQVGFRRKGLRML 297
 Db 185 SLMSTPPQYQYEQKRRSLIDM-----KTRPQSGGQLPAAKQGG-FQLSLTGLQLT 237
 OY 298 PPAISARLSKLLKLSWLSSTKSEKGYHLTYETREGVVSLOSRSIVMTVPSYVANSIL 357
 Db 238 VERIEKQI--KLTKYKKGKTKVRLSHSNGSLSYLEDLNG-VTLDADSVYTPARHKAAGML 294

OY 358 RPLSVAADALSNFYYPVAVGATVTSYRQEARLDERLVDGLKGFQGLHRTGQVETLGTI 417
 Db 295 SEL--PAISHLTKMHSTVAANVALGPEEGSYGMH-----EGTGFVLSRNSDPAITACT 346
 OY 418 YSSSLPNNAPRGRVILLNYIGAKRPELLSTESOLVYVNDRLKMLIKRAODPLVY 477
 Db 347 WNNRKKRNHAPRSGKTLRLAYVYGRKAGDESYVDSDDIINIVLEDKKVV--NINGEPEMT 404
 OY 478 GYRVWFOALPQFLVGHLDLSTAKAAMNDNGLELFLGGVNYVYVYVYVYVYVYVYVYVY 537
 Db 405 CYTRHESHPQYHVGHKKIKELREAL-ASATYRQYVMYMGASPEVGIPTCDIGKAAVSD 463
 OY 538 YVGFEL 542
 Db 464 ALTYL 468

RESULT 6
 PROX_PROFR STANDARD; PRT; 527 AA.
 AC 032434;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
 GN HEMX.
 OS Protoporphyrin freudenreichii shermanii.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Propionibacterineae; Propionibacteriaceae;
 OC Propionibacterium.

OX NCBI_TaxID=1752;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IFO 12424;
 RX MEDLINE-97306686; PubMed-9163953;
 RA Hashimoto Y., Yamashita Y., Murooka Y.;
 RT "The Protoporphyrin freudenreichii hemYHBKRL gene cluster, which
 RT encodes enzymes and a regulator involved in the biosynthetic pathway
 RT from glutamate to protoheme";
 RL Appl. Microbiol. Biotechnol. 47:385-392(1997).
 CC -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 CC IX TO FORM PROTOPORPHYRIN IX.
 CC -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2).
 CC -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
 CC -!- PATHWAY: PENTULIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOSOL; PLASMA MEMBRANE (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D85417; BAA21909.1; -
 DR InterPro; IPR000171; -
 DR Pfam; PF02032; Phytoene_dh; 1.
 KW Prophylin biosynthesis; Heme biosynthesis; Oxidoreductase;
 FT NP_BIND 33 38 FAD (POTENTIAL).
 SO SEQUENCE 527 AA; 55267 MW; 579CG0A3318BF938 CRC64;

Query Match 13.2%; Score 375.5; DB 1; Length 527;
 Best Local Similarity 27.2%; Pred. No. 7.1e-21;
 Matches 144; Conservative 94; Mismatches 210; Indels 81; Gaps 22;
 OY 58 TVPSSAVDQ----GP-AAELDCVYVAGISGLCIAQVMSANPRLMTVEARDRAGNIT 112
 Db 8 TTPPTVYSGTAPRPSHCHLVVYGGITGIAAAMQMGARVYVEESDDHFGKQVVT 67
 OY 113 VERDGYLMEEGPNSFQPSD-MLTMAYDCGLKBDLVLDGPNAPRVIY-----KGLRIPSK 169
 Db 68 DRDDEFVLEGGPDSFVAVRPAALKLIBELGLSDQVIA--PEGGRVSLSRGLRPFMPAG 125
 OY 170 LTDL-----PFF--DLMSIPKLRAGFGPIGLRPSRPGHEESYQVFRNMLGSEVFBR 221
 Db 126 MGVMLPTRMWPFVYTVTLVSWPDKIRAGLDLY-IPRRLPDHVAIAGVRLRDLGDIYRRF 184
 OY 222 IEPCCSVY 274
 Db 185 ADPMVGGIYVAGIDELSDAVLPSLRDNERHSIMVAVSILAGFASRARRARORAAQONNAQ 244
 OY 275 RDPRLPRKKGQ-----TYG-----SFRKGLRMLPDAISARL-----GSKLKLKLSKLS 319
 Db 245 QNSSHQNSTGQNSAGTGPAPRFLRGLGLDLDALVDQLAAGVLELVN--TSVD 301
 OY 320 KSEKGYHLTYETREGVVSLOSRSIVMTVPSYVANSILRPLSVAADALSNFYYPVAVGAV 379
 Db 302 LLGRDGYHLS----DGRV-LPADAVLVLAGVYASSARLRLRQLRAARAAALQIPLASTTV 356
 OY 380 TITYPQEARLDERLVDGLKGFQGLHRTGQV-----VELLGTIYSSSLPNNAPRGRV 433
 Db 357 SLAMPVSAF-----DVAEPDSQGWLEPADAGVPSGLTASSIKRGAAPDGSVYL 402
 OY 434 LLNVIYGAKNPELLSTESOLVYVNDRLKMLIKRAODPLVYVYVYVYVYVYVYVYVYVY 493
 Db 403 MRVEVPPDKRP-LTDAPDDELLSAVIDHVRPL--GVHGEGLQLTRMKRKMVKYTIYGH 459

OY 494 LDTLSTAKAAMNDGL----EGLFLGGRVYSGVALGRCVGAAYEVA5EV 538
 Db 460 LE-----RAAVVDSTLAEQRPTWAVASALHGVLPCICIDBARHSADEV 503

RESULT 7
 PROX HUMAN STANDARD; PRT; 477 AA.
 AC P50336: 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
 GN PPOX.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiinda; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=95292621; PubMed=7713909;
 RA Nishimura K., Taketani S., Imokuchi H.;
 RT Cloning of a human cDNA for protoporphyrinogen oxidase by
 RT complementation in vivo of a hemg mutant of Escherichia coli.;
 RT J. Biol. Chem. 270:8076-8080(1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=96367087; PubMed=8771201;
 RA Hall T.A., Dalley H.A.;
 RT "Human protoporphyrinogen oxidase: expression, purification, and
 RT characterization of the cloned enzyme.";
 RT Protein Sci. 5:98-105(1996).
 RL [3]
 RP VARIANT VP ARG-232, AND VARIANT HIS-304.
 RX MEDLINE=97005368; PubMed=8852667;
 RA Dezbach J.-C., Puy H., Robreau A.-M., Lamoril J., da Silva V.,
 RA Grandchamp B., Nordmann Y.;
 RT "Mutations in the protoporphyrinogen oxidase gene in patients with
 RT variegate porphyria.";
 RT Hum. Mol. Genet. 5:407-410(1996).
 RL [4]
 RP VARIANT VP TRP-59, AND VARIANT CYS-168.
 RX MEDLINE=96241580; PubMed=8673113;
 RA Meisner P.N., Dalley T.A., Hift R.J., Zlman M., Corrigall A.V.,
 RA Meisner P.N., Dalley T.A., Hift R.J., Zlman M., Corrigall A.V.,
 RA "A R59W mutation in human protoporphyrinogen oxidase results in
 RT decreased enzyme activity and is prevalent in South Africans with
 RT variegate porphyria.";
 RT Nat. Genet. 13:95-97(1996).
 RP VARIANT VP CYS-152.
 RX MEDLINE=98434227; PubMed=9763307;
 RA Frank J., Pol-Fitzpatrick M.B., King L.E. Jr., Cristiano A.M.;
 RT "The genetic basis of 'Scarsdale Goumet Diet' variegate porphyria: a
 RT missense mutation in the protoporphyrinogen oxidase gene.";
 RT Arch. Dermatol. Res. 290:441-445(1998).
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 CC -1- TO FORM PROTOPORPHRIN IX.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2).
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER.
 CC -1- PATHWAY: PERULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
 CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,
 CC LIVER, SKELETAL MUSCLE, KIDNEY, AND PANCREAS.
 CC -1- DISEASE: DEFECTS IN PROX ARE THE CAUSE OF PORPHYRIA VARIEGATA
 CC (VP). A DISEASE CHARACTERIZED BY SKIN HYPERPIGMENTATION AND HAIR
 CC HYPERTRICHOSIS, ASSOCIATED WITH ACUTE ATTACKS, LIKE THOSE OF ACUTE
 CC INTERMITTENT PORPHYRIA.

CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as their content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).

DR EMBL: D38537; BAA07538.1;
 DR EMBL: U26446; AAA67690.1;
 DR MIM: 600923;
 DR MIM: 176200;
 KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
 KW Flavoprotein; FAD; Mitochondrion; Disease mutation; Polymorphism.
 FT NP_BIND 9 14 FAD (POTENTIAL).
 FT VARIANT 59 59 R->W (IN VP).
 FT VARIANT 152 152 /FTID=VAR_003686.
 FT VARIANT 168 168 /FTID=VAR_003687.
 FT VARIANT 168 168 R->C (IN VP).
 FT VARIANT 232 232 /FTID=VAR_003688.
 FT VARIANT 232 232 G->R (IN VP).
 FT VARIANT 304 304 /FTID=VAR_003689.
 FT VARIANT 304 304 R->H.
 FT FTID=VAR_003690.
 SQ SEQUENCE 477 AA; 50765 MW; 2444DEAC2E6C33EE CRC64.

Query Match 12.8%; Score 363.5; DB 1; Length 477;
 Best Local Similarity 25.5%; Pred. No. 5e-20;
 Matches 138; Conservative 83; Mismatches 170; Indels 151; Gaps 23;

OY 75 VYVGAGISGLCLTAQVMSANV-----PMLAVTARDRAGCNITTV-RCGYLMEBEPN 125
 Db 6 VYVGAGISGL-----AASYHLRRAPCRPKVVYVSSESLGGMWMSVKGPNCAITELBPR 59
 OY 126 SFOPSDPM-----LTMAVDDGLKDDI--VLDGPNAP--RFVLMKGLRVPYPSKLTPLDPPFD 177
 Db 60 GTRPAGALGARRILLVYSELGDLSEVLPVPGDHPAQNRLVLYVGGALHALPGLRBL 115
 OY 178 LMSIPGKLRAGFGPIGLRPSPP-----GHE--BSVQVPRNRNIGGEVF 218
 Db 116 -----LRSPSPSKPLFWAGLRELTFRKGRKPRDPTVHSFAQRRLGPEVA 159
 OY 219 ERLEIFPCGGYVVGDSKLSMKAARFQKWKLEPTGGSTTGFRKIKRSTPRKARPR 278
 Db 160 SLAMDLSLCKGVAFAGNSRELSTIRSCPSLRFQAEOTHRSTLLGLLGA-----GR 207
 OY 279 LRPKRGQTV-----GSPFKGLRMLPDAISRALGSK-----LKLS----- 312
 Db 208 TQPDPSALLRQALAEBSWSLSLRSGLGEMLRQALFHHLHLSRGSVSLVRGQPVGCLSLQAEGR 267
 OY 313 WKLSITTKSEKGGYHLLTETPBGVVSLSQSRSLVMTPVPVSVASNIIRPLSVAAADALSIFY 372
 Db 268 WKY-SLRDS-----SLEADHVISAIIPASVSELPLPAPAAPLARALSMT 310
 OY 373 YRPVGVVTTISYQDEAIRDRLVVDGLKFGQIHPRTQGVETLGTIYSSLSLPRNR--ARKG 430
 Db 311 AVSVAVAVVNIQYQGAH-----PVGQGFGLVPSSEDDPGVGLIYDVAAPEDQGSPPG 362
 OY 431 -RVLLANTYTGAKNPEILSKTESQLEVVVDRLRMLIKPKAKDDPLV--VGVRVWP----- 483
 Db 363 LRYTVW--LGG-----SMQLTLPASGCVLSDELQQRQAEAAALQDLGKKEPMSCL 411
 OY 484 ----QAIPLVGHLDLTLSTAKAAMNDGLEGLFGLGNYVSGVALGRCVGAAYEVA5EV 538
 Db 412 VHLKKNKCIQYTLGHMOKLESARQELTLHRPLTLAASV--EGVAVNDICISGRQAASV 470
 OY 539 TG 540
 Db 471 LG 472

RESULT 8
 PPOX_MOUSE STANDARD; PRT: 477 AA.
 AC P51175; P97344;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
 GN PPOX.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96132666; PubMed=8554330;
 RA Dalley T.A., Dalley H.A., Weissner P., Prasad A.R.;
 RT "Cloning, sequence, and expression of mouse protoporphyrinogen
 RT oxidase."
 RL Arch. Biochem. Biophys. 324:379-384(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Erythrocyte;
 RA MEDLINE=95331315; PubMed=7607249;
 RX Takekani S., Yoshinaga T., Furukawa T., Kohno H., Tokunaga R.,
 RA Nishimura K., Inokuchi H.;
 RT "Induction of terminal enzymes for heme biosynthesis during
 RT differentiation of mouse erythrocyte cells."
 RL Eur. J. Biochem. 230:765-765(1995).
 RN [3]
 RP SUBCELLULAR LOCATION.
 RC TISSUE=Liver;
 RA MEDLINE=88153682; PubMed=3346226;
 RX Ferrreira G.C., Andrew T.L., Kerr S.W., Dalley H.A.;
 RT "Organization of the terminal two enzymes of the heme biosynthetic
 RT pathway. Orientation of protoporphyrinogen oxidase and evidence for a
 RT membrane complex."
 RL J. Biol. Chem. 263:3835-3839(1988).
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 CC IX TO FORM PROTOPORPHYRIN IX.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2).
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
 CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.
 CC -1- INDUCTION: DURING ERYTHROID DIFFERENTIATION.
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 DR EMBL: U25114; AAA96003.1; -;
 DR EMBL: D45185; AAA08126.1; -;
 DR MGD: MG1:104968; PPOX.
 KW Porphyrin Biosynthesis; Heme biosynthesis; Oxidoreductase;
 KW Flavoprotein; FAD; Mitochondrion.
 FT NP_BIND 9 14 FAD (POTENTIAL).
 FT CONFLICT 64 64 A->T (IN REF. 2).
 FT CONFLICT 66 66 A->P (IN REF. 2).
 FT CONFLICT 108 108 L->S (IN REF. 2).
 FT CONFLICT 427 427 W->C (IN REF. 2).
 FT SEQUENCE 477 AA; 50870 MW; 8CFB48120729BDE6F CRC64;

Query Match 12.7%; Score 362.5; DB 1; Length 477;
 Best Local Similarity 24.8%; Pred. No. 5.9e-20;
 Matches 133; Conservative 76; Mismatches 188; Indels 139; Gaps 19;

QY 75 VIVGAGISGLCTAQ-VMANPNLMTVEARDRAGNTTYVE-RDGYIWMEDGNSPQSD 131
 :::::::::::::::::::::
 DB 6 IVGGISGLIASYMLRGPSPKYLIVGSKRLGWTIRSIGSDAIFELGRGIRPAG 65
 132 PM-:-LTMADVCGKDKD-LVGDPMNP-RFVLMKGRIVPYSKLDLPDFLMSIRG 183
 66 ALGARILLIVSELGLESEVLPVRGHPAQNRFLLYVGTFHPPLSGLSL-:-: 115
 QY 184 KLRAGFGPIGLRSPSP-:-:-----GHE--ESVEQFVRRNIDGSEVFRLEIEP 224
 116 -----LRRPSPFSPKPLFWAGIRELKRGEPEPEYHSHFQRRLGPEVSIANDS 165
 QY 225 FSGVYVDPKSLSKMAAFGWKLEFNGGSTIGTFPAIKERSSTPRAPDPLPKPKG 284
 166 LRGVFAVGNRSRLESTRSCPSPSLFQAETHRSTLLGLIGAGOS-----POPDS 213
 QY 285 QTV-----GSPFRGLRMLPDATISARLGSK-----TKLS---WKLSSI 318
 214 SLIRQARARRMWSQLRGGLEVLQALHNLHLSKGVTVLVSGQPVCGLSLQDPERKRKVSLL- 272
 QY 319 FSKERGVNHLIYETPEGVVSLQSRSIYMTVPVSYVASNLIIRPLSYAAADALSNFYYPVGA 378
 273 -----GQSLEADHIISAIIVASELSKLLPBAAPLANILTITAVSVAV 316
 QY 379 VYISVPOAIRPDELVDGELKFGQLHRRQGVETPLGTYSSLSLFPNR---APKGRVILL 435
 317 VMLQVRGACL-----PVQGFHILVPSSEDPVTLGIYVDSVAFPEDDGNPDLRIVYM 368
 QY 436 NYIGAKNPELITKSTESQVEYDRDLRKMLIKPKAOD-----PLVVGVRVWPQ 484
 369 --LGG-----YWDLKRAAGQLSEPELFQDQAQDAANAATQLGKRPESHCLVHDKN 417
 QY 485 ALPQFLVGHLDLTLSTAKAAMDNGLEGFLGNGVYVAVLGRVCGEAVYVASEYTVG 540
 418 CIPQYTIQHWQKLDLSAMQFLFAQRRLPFLTAGASY-EGVAVNDGIESGGQAQAVAVLAG 472
 DB

RESULT 9
 YAM7_SCHPO STANDARD; PRT: 490 AA.
 AC Q10062;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPOPHETICAL 53.5 KDA PROTEIN CLP5.07C IN CHROMOSOME I.
 GN SPAC1P5.07C.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=977;
 RA Genles S., Odell C., Churcher C.M., Barrell B.G., Rajandream M.A.,
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) AND TO
 CC YEAST YER014W.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 DR EMBL: Z68136; CAA92235.1; -;

KW Hypothetical protein. 490 AA; 53461 MW; CB679F6A437547AA CRC64;

Query Match 10.6%; Score 302.5; DB 1; Length 490; Best Local Similarity 22.8%; Pred. No. 2.1e-15; Matches 120; Conservative 103; Mismatches 213; Indels 87; Gaps 20;

72 LDCVYVAGISGLCIAGVMSANYPNLMT--EARDRAGNITVVE-----RDGYLMEB 122
1 MSALICGGGIAGLSTAVYLAIRLIPKCTIDLYEKGPRLGWLQSVKIKPCADSPRTGVLFRQ 60
123 GPNSEFQSD-----PMLTMAVDCGKDDLDLVGDPNAP-----RFLVWKGKLRVPSKLTLDLP 174
61 GPRFLRPAVAGLANLNDLISKLGEDKLLNISSNSPAAKRRYYIDRLEIYS-----114
175 FFDLMSIPGKLRAGFGPIGDRPSP-----PGHSESEVDFVRRNLGGEVEERL 221
115 -----SILGSIKSIQMP-ALRPPMLAMMLEPFRKSKRSDTDESEVSGEMRRRFRGKVTDRY 168
222 IEPFGCGVYVGDPSKLSMKA--RQKWKLEBTGSGITIGFKAIKER---SSTPKAPRDP 277
169 MSAMINIGIYAGDLDLNDLSMSSMFGFLAKIKKYGNTLGLIRALLARREILSPAEKALKAA 228
278 RLPRP-----KQOTVSGFRKGLRMLPDAISARLGSKIKLTKLMSITRSEKGGYHL 328
229 LLAEPKTAELSNMKSMTSMARFKEGITITLSTLSTADELKKRPNVKIHLNKRKAKT-----L 282
329 TYETPEGVSVLSQKRSIYMTVPSYVANSILRPLSVAADALSINFRYPPVGAVTISYQEA 388
283 VPHKQGLVYDNGQAYEYVAVANSRMLLENLISCPKMEPTPTSSVY-----VVNYVY----- 333
389 RDELVDVGELEKGFQGLRPR---IQGVETLGTIYSSSL-PFNRRAPKGRVILLNNTYGA--- 441
334 KDPNVL--PIRGEFLIIPSCPTNPNPHVLDIVEDSEGNENSGSKVTVMK-----GGAYTP 387
442 KNPIILSKTESQLEVEVDRDLRKM-LIKPKAODPLVYGVHVPQAIPOFLVGHLDLTLSTA 500
388 KNISLIPIPNPEAVNNMALKLQHLTKLSSK---PTLNLNATLQONCIPQYVAVGHQDNINSL 444
501 KAAMNDNGLEGLFLGNYVSGVALGRCVEGAYEYVASEVYTGPLS 543
445 KSMIEKMMGGRILLTGSWYNGVSGIDCIMGHSPARKLASLMM 487

FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRIN GEN IX TO FORM PROTOPORPHYRIN IX.
- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) - PROTOPORPHYRIN-IX + H(2)O(2).
- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
- SUBUNIT: HOMODIMER (BY SIMILARITY).
- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
- SIMILARITY: BELONGS TO THE PROTOPORPHYRIN GEN OXIDASE FAMILY.
- CAUTION: SEEMS SHORTER IN THE C-TERMINAL THAN CORRESPONDING HEMOIN OTHER BACTERIA.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.1sb-sdb.ch/announce/ or send an email to license@1sb-sdb.ch).
EMBL: AL021186; CAAL16020.1;
Tuberculin; Rv2677c;
Porphyrin biosynthesis; Oxidoreductase; Flavoprotein; FAD;
Heme biosynthesis.
NP_BIND 10
SEQUENCE 426 AA; 44382 MW; 221952B3977D3B0 CRC64;

Query Match 7.8%; Score 221; DB 1; Length 426; Best Local Similarity 22.8%; Pred. No. 2.5e-09; Matches 108; Conservative 74; Mismatches 210; Indels 82; Gaps 17;
76 IYVAGISGLCIQ-OVMSANYPNLMT--EARDRAGNITVVERDGYLMEGPNSEFQSDP 132
8 VVGGISGLTSAIRLQAVGDATITTEPRADRLGGLVLRTEHIGGQPPDLDAEAFVLRPP 67
133 -MLTMAVDCGLKDDLVLDGPNAPRFLVWKGKLRVPSKLTLDLPEFDLMSIPGKLRAGCP 191
68 EMPALLAELGL-SDRQLASTGAPRLIYSGQRRLHPPD-----TVVGIPTSSAGSMAGL 119
192 I-----GLRPS--PGHSESEVDFVRRNLGGEVEERLIEPFGCGVYVGDPSKLS 238
120 VDDATLARIADAARPEFTWQVSDPVAVDLVDPRGDDVAVARSVPLLSGYAAGSAATIG 179
239 MKAAFGKWKLEBTGSGITIGFKAIKERSYKAPRDRPKPKGQTVGSEFRKGLRMLP 298
180 LRAAASVAALDRGA-----TSYTDVAVRQALPFGSGGYPVFGALDGGYVLL 226
299 DAISARLGSKLSKLSITRSEKGYHLIYETPEGVVSIQSRSIWTVPSYVANSILR 358
227 DGLVRR-----SHVHWVRAVYVODLER-GWVLRDETEG---RWDADVIVILAVPAPRLARLVD 278
359 PLSVAADALSINFRYPPVGAVTISYQEAIRDERLVDGELKFGQLHPRTOGVEVTLG--- 415
279 GIAPRTHAARQIVSASSAVVALAVPG-----GTAFFHCSGVYLVAGDES 322
416 -----TIYSSSLFPRARKGRVILLNNTYGAKNPE-ILSKTESQLEVEVDRDLRKMLIK 468
323 PRAKATITLSSRWK---GQRGDVALLRLSFGFRGEPALTAASDDQLLMAAADDL--VYVF 376
469 PRAQODPLVYGVRRVWQAIPOFLVGHLDLTLSTAKAAMNDNGLEGLFLGNYVSGV 522
377 GVAVDVAVRVRWLEAMPQYGRHADVVAELRAGLIPPT---LAVAGSYLDGI 426

RESULT 11
PROX_MYCTU STANDARD; PRT; 451 AA.
AC 050008;
15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).

GN HEMV.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRIN
 CC IX TO FORM PROTOPORPHYRIN IX.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2).
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U15181; AAA62958.1;
 KW Porphyryn biosynthesis.
 KW Heme biosynthesis.
 FT NP_BIND 10
 FT SEQUENCE 451 AA; 46580 MW; D76BE1655CA2056 CRC64;

Query Match 7.6%; Score 216.5; DB 1; Length 451;
 Best Local Similarity 20.6%; Pred. No. 5.8e-09;
 Matches 105; Conservative 79; Mismatches 193; Indels 133; Gaps 18;

QY 76 IVGAGISGLCTA---QVMSANRYNMLTQEARAGNITTVYERDGYLWEEGNSF---QP 129
 Db 8 VVGGISGLTAAAYRLVAATGDDVAITLTFPDRGLGVLRTFCVGGQPMDDGAEAFLLRPP 67
 QY 130 SDPMLTMAVDDGLKD-----DVLGDPNAPRYLWKGKLRPV 166
 Db 68 EYPAL- LAELGISERQATTDARPLIYSQORLHSLPDDVAG-----I 109
 QY 167 PSKLTDLPEFDLMSIPGKLR---AGFGPGLRPS--PGHEESVEQFVRRNLGGEVEER 220
 Db 110 PSSAT-----SVAGLVDDATVVARIGAEVAVRPLSWEPGSDPMAELVADRREGEQAVAR 161
 QY 221 LIPEFGGVYVGDPSKLSMKRAAFKRWKLETTGGSTIGGTFKAIKERSSPPKAPRDRPLP 280
 Db 162 LVDDLGGVAVGSAATGTLRGAPASVAALDCCG-----TSLMEAVRQGLPP 208
 QY 281 KPRGQVGFSEFKGRMLPDAISARLGSKTLNKLSTITSEKSG--GYHLVYETPEGVYS 338
 Db 209 VAAPRPGVAGLDGGVQVYLDELVRR-----SRLOVAAATVVGIDRGTGWTLVYDDG---AC 261
 QY 339 LOSRSITVWPVSYAVSNTLRLPLSVAAADALSNFYPPVGAVTISYQEARLDRERIVDDEL 398
 Db 262 WSAAGVITAVPAPRLVRLLOOIAIPRVAAASRIYSASSAVVALSVPRDPTFPQN--SGVL 319
 QY 399 KGFQO-----LHPRTOGVENTLGTIYSSSLPNNRPAKRVLLIANTIG-----A 441
 Db 320 VASGERLRARAAVITSSRKKWGLQ-----GDTQLVRLISFKKFGDDOVASTA 362
 QY 442 KNPEILSKTESQLEVVVDRDLRKMILKPKAODPLVVGVRWVPAIIPQFLVGHLLDTLSTAK 501
 Db 363 SDDLFLMAVAVSILAAYVDVTV-----DPVDVQVQKWIADAMPQYGGHGHADVAVR 412
 QY 502 AANNMNDLBEGLFLGGNYSVVALGRCEGA 531
 Db 413 AGLPPT---LVVAGSHMDGIGVPAACISAA 438

RESULT 12
 PROX_YEAST
 ID PROX_YEAST STANDARD; PRT; 539 AA.
 AC P40012;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
 GN HEM14 OR YER014W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96224138; PubMed=8621563;
 RA Camadro J.-M., Labbe P.;
 RT Cloning and characterization of the yeast HEM14 gene coding for
 RT protoporphyrinogen oxidase, the molecular target of diphenyl
 RT ether-type herbicides.
 RL J. Biol. Chem. 271:9120-9128(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berro A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SUBCELLULAR LOCATION.
 RC STRAIN=ATCC 25657 / D273-10B;
 RX MEDLINE=95096046; PubMed=7798202;
 RA Camadro J.-M., Thome F., Brouillet N., Labbe P.;
 RT Purification and properties of protoporphyrinogen oxidase from the
 RT yeast Saccharomyces cerevisiae. Mitochondrial location and evidence
 RT for a precursor form of the protein.
 RL J. Biol. Chem. 269:32085-32091(1994).
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 CC IX TO FORM PROTOPORPHYRIN IX. OPTIMUM ACTIVITY IS OBTAINED AT PH
 CC 7.2.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2).
 CC -1- COFACTOR: FAD; CONTAINS 1 MOLE OF FAD PER MOLE OF ENZYME.
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
 CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z71381; CAA95981.1;
 DR EMBL: U18778; AAB64547.1;
 DR SGD: S00008116; HEM14.
 KW Prophyryn biosynthesis; Heme biosynthesis; Oxidoreductases;
 KW Flavoprotein; FAD; Mitochondrion.
 FT NP_BIND 18
 FT MOTIF 422
 FT MOTIF 424
 FT MOTIF 424
 FT SEQUENCE 539 AA; 59703 MW; 5ECCBDD1C033BA5B1 CRC64;

```

Query Match      6.6%; Score 189; DB 1; Length 539;
Best Local Similarity 20.3%; Pred. No. 9e-07;
Matches 111; Conservative 95; Mismatches 202; Indels 138; Gaps 24;

QY 68 PAALDCVIVGAGISGICIAOVMSANYPNLMVY--EARDRAGG-----NITVEERDGYLM 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 10 PRAKV--AVYGGVSGICFYFPLSKLRPVEIILFESQNMGTGWIYSCWPRDMSGPRIML 67
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 EEGPNSFQ-PSD-----PMLTMAVDCG-LKDDLVLDGPPNAPRVLYMKGKLR 164
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 68 EKGPRLTRGVSDDGTVLIMDTLKLKDKREAVYQSIQSDKGCIDAKRFLDPPS-----DKLIV 119
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 165 PVPSKTLDELPEPDMST-PEGLKLRAGGRPIGLRSP-PGHESEVEQPVRRNLGGE-VEERL 221
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 120 QVPKSTITTKFLNPLGKLLTGMMGEMFRKRSPPHGQDESEVSCDRFRFGNNTYSNMW 179
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 222 IEPFCSGYVVDPSKLSMKAAPGKVMKLEBTEGGSIIGTDFPKAIRKSSSPPKAPR----- 275
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 180 ISALNRGITYGDDVYLSLAKRTPFKKIYVNLKHGSMNTQAMIDNMRGKRSRKRTEMLHQSLT 239
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 276 -----DPLRPPKPGQTV--GSFRK-----GLRMLPDALSAKLGs----KTKLSMK 314
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 240 GCLNDYSNAFGKDRSKLDLSDNLTAKRPMILGAGLETFEPKIVRNALNERNKYKIVTGNP 299
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 315 LLSI-----PKSEKGGYHLIYE-----TPEGVYSL--QSRSIYMTVPVSYVSN 355
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 300 VTQIMKRRPANEETITGLAKAGSDQETFDHU.RLITTPKIKLKPDKONSISKLIDEIQSN 359
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 356 ILRPLSYAADAALSNFYYPVGAVTTISYPOEATIRDERLYVDEIKPFGQLHPRP--OGVER 413
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 360 TI-----ILVNYLTP-----NKKDVIDADLQGGFYLVKSNKNGKL 395
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 414 LGTITSSYLEPPNRAKGRVLLNLTIGGAKKPEILSK-----TESQLVEYVDR 460
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 396 LGVTFEDSVIERNEFKP-----LEFDKL--STPNPMLNKYTKVYAMIGCCMLNHEHGPVPSR 448
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 461 DLRKMLLKPKRQDPLVYGVYR-----VMPQATIPQFVYGHLDLSTAKKAMNDNGLE 510
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 449 EVTINAAVDALNHLGISNKKDLEAGWEPFIADRCLEPRFHVGDADWQERBARLQESYQG 508
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 511 GLPFGG 516
      | | | | |
DB 509 TVSYGG 514

RESULT 13
AOFH_MYCTU  STANDARD;  PRT;  448  AA.
ID  AOFH_MYCTU  STANDARD;  PRT;  448  AA.
AC  053320;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  PUTATIVE FLAVIN-CONTAINING MONOAMINE OXIDASE RV3170 (EC 1.4.3.-).
GN  RV3170 OR MV014.14.
OS  Mycobacterium tuberculosis.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CC  Actinomycetales; Corynebacteriales; Mycobacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1773;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=98295987; PubMed=9634230;
RX  MBLINP=98295987;
RC  STRAIN=H37RV;
RA  Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA  Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA  Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA  Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holtroyd S.,
RA  Oliver S., Jagsen K., Krogh A., McLean J., Moule S., Murphy L.,
RA  Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA  Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA  Taylor K., Whitehead S., Barrrell B.G.;
RA  "Deciphering the biology of Mycobacterium tuberculosis from the
RA  complete genome sequence".
RL  Nature 393:537-544(1998).

```

```

CC -1 COFACTOR: FAD (POTENTIAL).
CC -1 SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL station on
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
-----
CC EMBL; AL021646; CAAT1635.1;
DR TubercuList; RV3170;
DR InterPro; IPR002937;
DR Pfam; PF01593; amino_oxidase_1.
KW Hypothetical protein: Oxidoreductase: Flavoprotein; FAD.
FT NP_BIND 8 63
SQ SEQUENCE 448 AA: 48439 MW; 93FAZC74B3F13D7A CRC64;

```

Query Match 5.0%; Score 141; DB 1; Length 448;

Best Local Similarity 21.6%; Pred. No. 0.0029;

Matches 65; Conservative 59; Mismatches 119; Indels 58; Gaps 13;

```

QY 68 PAALDCVIVGAGISGICIAOVMSANYPNLMVTEARDRAGGNITTVERDGYLMGEPNSF 127
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 4 PPTWYDYYVYVGGAFGLAAARELVRQGHVLYFEGRDHVGGRSLTGRVAAGVPADMGGSFI 63
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 128 QPS-DPMLTMAVDCGLKDDLVLDGPNAP-----RFLMKKLRVPSKLDLPEPDM 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 64 GPQDAVALALATE-----LGIPTTPHRDGRNVYWRSSARSRKRTIKRLSITGLID 115
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 IPGKLKRGFRPT--GLRSPRPGHESEVDEQVFRNLGGEVEFLIEPFCSGYVVDPSKLS 238
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 116 I-GRLEWQFERLARGVPPAAPADARARELDDYSIG--EMLRIVR-----ATSSSRN 164
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 239 MKRAAFGKV-WKLEMTGGSIT-----GTFKAIKERSSTPPKARPDLKPKGQTV 288
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 165 IMAIMRVYTWGCEPDVSMILHARARYVRAAGGLRLLDVKNGA----QQDRVP----- 212
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 289 SPRKGLRMLPDAISRLGSKLTKLMSKLSITKSEKGGYHLIYEPREGVYVLSQSRIVMTV 348
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 213 --GGTQOIAQAAAADLQARVILLN--AAVRRIDRHGAGVTVSDQG--QAEAGCVIVAI 264
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 349 P 349
      |
DB 265 P 265

RESULT 14
CRTI_NARPS  STANDARD;  PRT;  570  AA.
ID  CRTI_NARPS  STANDARD;  PRT;  570  AA.
AC  040406;
DT  01-OCT-2000 (Rel. 40, Created)
DT  01-OCT-2000 (Rel. 40, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)
DE  PHYTOENE DEHYDROGENASE PRECURSOR (EC 1.3.-.-) (PHYTOENE DESATURASE).
GN  PDS1 OR PDS.
OS  Narcissus pseudonarcissus (Daffodil).
OC  Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
CC  Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae; Narcissus.
OX  NCBI_TaxID=39639;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  TISSUE=Paracolla;
RC  TISSUE=Paracolla;
RA  Al-Badili S., Beyer P.;
RA  "A cDNA encoding phytoene desaturase from daffodil.";
RA  (In) Plant Gene Register PCR95-131.
RN  [2]
RP  SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA  TISSUE=Paracolla;
RC  MEDLINE=96237452; PubMed=8653112;
RA  Al-Badili S., von lintig J., Haudruck H., Beyer P.;

```

RT "A novel, soluble form of phytoene desaturase from Narcissus
 pseudonarcissus chromoplasts is Hsp70-complexed and competent for
 flavinylation, membrane association and enzymatic activation.";
 RT Plant J. 9:601-612(1996).
 RL [3]
 RN SUBCELLULAR LOCATION.
 RP MEDLINE=97433278; PubMed=9288918;
 RX Honk M., Hoffmann B., von Lintig J., Schledz M., Al-Badilli S.,
 RA Hobeika E., Klehidig H., Beyer P.;
 RA Chloroplast import of four carotenoid biosynthetic enzymes in vitro
 RT reveals differential fates prior to membrane binding and oligomeric
 RT assembly";
 RL Eur. J. Biochem. 247:942-950(1997).
 CC -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
 CC INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
 CC DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
 CC -1- COFACTOR: FAD (PROBABLE).
 CC -1- PATHWAY: CAROTENOID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST (CHROMOPLAST). EXISTS AS AN
 CC INACTIVE SOLUBLE FORM AND AN ACTIVE MEMBRANE-BOUND FORM.
 CC -1- TISSUE SPECIFICITY: EXPRESSED MORE STRONGLY IN FLOWERS THAN IN
 CC LEAVES.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION INCREASES DURING FLOWER
 CC DEVELOPMENT.
 CC -1- SIMILARITY: TO OTHER PLANTS OR CYANOBACTERIAL PHYTOENE
 CC DESATURASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X78815; GAA55392.1; -
 DR Mendel; 13184; RARPS; P0S1; 13184.
 DR InterPro; IPR002025; -
 DR Pfam; PF01593; Amino_oxidase; 1.
 DR Carotenoid biosynthesis; Oxidoreductase; FAD; Chloroplast; Membrane;
 KM TRANSIT peptide.
 FT CHAIN 92 570 CHLOROPLAST (PROTEIN).
 FT NP_BIND 104 120 PHYTOENE DEHYDROGENASE.
 FT SEQUENCE 570 AA; 63790 MW; 654F569F5B83BE77 CIRC64;

Query Match 4.9%; Score 139; DB 1; Length 570;
 Best Local Similarity 21.1%; Pred. No. 0.0058;
 Matches 111; Conservative 92; Mismatches 178; Indels 146; Gaps 24;

DB 351 LGNGNITIGDAVY-----VAA-----PVDILKLLDPMRBPYFKLD-K 390
 OY 398 LKGFGLHPRTOGVETLGIYSSLF-----PNRAPKGRVLLN 436
 DB 391 LVGVPIINWHIMFDRKLTNTYDHLFTSPSPLLSYAADMSTYCKREYDNNSS----MELLY 446
 OY 437 YIGGAKNPILLKTESQVLEVVDRDLRMLIKPKAQDPLVGVVHPQALPQLVGHLDI 496
 DB 447 F---AAAEWISRSSEIIEIRMKELAKIFPDEIYAD-----QSNKRIKIYHV-- 491
 OY 497 LSTAKA-----MNDNGLEGLFGLGNVYSGVALGRCEGA 531
 DB 492 VKTPRSYKTIPTDCEPCRPLOQSPTEGFLAODITNOKYLA-SMEGA 537

RESULT 15
 AOF_ONCMY STANDARD; PRT; 522 AA.
 AC P49253;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 13-DEC-1998 (Rel. 37, Last annotation update)
 DE AMINE OXIDASE [FLAVIN-CONTAINING] (EC 1.4.3.4) (MONOAMINE OXIDASE)
 DE (MAO).
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Proteobacteria; Gammaproteobacteria; Sphingobacterii; Sphingomonadales;
 CC OX NBI_Ltaxid=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=95107295; PubMed=7808446;
 RA Chen K., Wu H.-F., Grimshy J., Shih J.C.;
 RT Cloning of a novel monoamine oxidase cDNA from trout liver.";
 RL Mol. Pharmacol. 46:1226-1233(1994).
 RN [2]
 RA IDENTIFICATION OF PROBABLE FRAMESHIFT.
 RA Balroch A.;
 RL Unpublished observations (JUL-1998).
 CC -1- FUNCTION: OXIDIZES BOTH 5-HYDROXYTRYPTAMINE (5-HT) AND BETA-
 CC PHENYLETHYLAMINE (PEA).
 CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
 CC H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN IN THAT, DUE TO A
 CC PROBABLE FRAMESHIFT, IT ENDS IN POSITION 499 AND LACKS THE C-
 CC TERMINAL REGION FOUND IN OTHER MEMBERS OF THIS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L37878; AAA64302.1; ALT_FRAME.
 DR InterPro; IPR001613; -
 DR Pfam; PF01593; Amino_oxidase; 1.
 DR PRINTS: PR00757; AMINEOXDASEF.
 KM Oxidoreductase; Flavoprotein; FAD; Transmembrane; Mitochondrion.
 FT NP_BIND 7 63
 FT BINDING 399 399 FAD (ADP PART) (POTENTIAL).
 FT TRANSMEM 493 513 POTENTIAL.
 FT SEQUENCE 522 AA; 58937 MW; F82D6D15364D646D CIRC64;

Query Match 4.7%; Score 134; DB 1; Length 522;

```

Best Local Similarity 22.0%; Pred. No. 0.012;
Matches 67; Conservative 48; Mismatches 119; Indels 70; Gaps 12;

QY 73 DCYIVAGISGICIAQVMSANYPNIMVTEARDRAGSNTTVERDGYLWEE-----GPNs 126
      |:::|111111|:::| | | | | | | | | | | | | | | | | | |
DB 8 DVIYVIGGIGLSAARLKEKGLSPVLEARDRVGGRTFTVONBQTKYVDLGGAYVGP- 66
      | : : | | : : | | : : | | : : | | : : | | : : | |
QY 127 FQSPDMLTMAVDCGLKDDLVLDPPNAPRFVLMKGIKLRPVPKLTDL--PF----- 175
      | : : | | : : | | : : | | : : | | : : | | : : | |
DB 67 ---QNRILRLAKKCGYKTKVNEEERLVHY--KGSYPFKGSFPDMNPFALMDYNNLM 121
      | : : | | : : | | : : | | : : | | : : | | : : | |
QY 176 --FDLM--SIPGKLRAGFGPIGLRPSPPGHESVVEQFVRRNIGEV----FERLIETPC 226
      | : : | | : : | | : : | | : : | | : : | | : : | |
DB 122 RKMDEMGSEIIPRE-----APWKAAPHAEMWDKMTMKQLDFDKICWTSARREFATLFV 171
      | : : | | : : | | : : | | : : | | : : | | : : | |
QY 227 SGYYVGDPSKLSMKAAFGKVKWKLEETGGSI-IGGFKAIKERSSTPKAPRDRPLPKKGO 285
      | : : | | : : | | : : | | : : | | : : | | : : | |
DB 172 NVNVVTSPPHEVS--ALMFLMIVVYKCCGGITRKIRISTTINGOQR----- 210
      | : : | | : : | | : : | | : : | | : : | | : : | |
QY 286 TVGSFRKGLRMLDPDAYSARLGSKLSLTKSSEKGGYHLYETPEGGVVSLOSRSIV 345
      | : : | | : : | | : : | | : : | | : : | | : : | |
DB 211 ---KPLGGSSQISQISEMAKKEISERVKME---SPVYKIDQTDGDMVEVELNKKEHYKAKYIV 264
      | : : | | : : | | : : | | : : | | : : | | : : | |
QY 346 MIVP 349
      | | |
DB 265 ATPP 268
      | | |

```

Search completed: July 3, 2001, 10:43:19
Job time: 612 sec


```

49 hhrGcSerValAlAlAspTyThrValProSerSerAlaValAsp
121 CAGATGTCCTCCGCAAGGATTAACAGTTCCTCCCGGAAGCGCAC
66 GtlyGlyProAlaAlaGluLeuAspCysValIleValAlaGlyTlIese 82
171 GGTAAATCAAGTCCCGGAGCTGGATGTGTGGATGGAGCGAATTAAG 220
82 rGlyLeuCysIleAlaGlnValMetSerAlaAsnTyProAsnLeuMetV 99
221 TGGACCTCGATTGCTAAGGTGATTCGGATTAATATCCCAATTTGATGG 270
99 aIhrrGluAlaArGAspArGAlaGlyGlyAsnIleThrThrValGluArG 115
271 TGACGGGAGGGAGGATCTGTCGGGGGAAACATAACGAGGTGGAAACA 320
116 AspGlyTyrlLeuTrpGluGluGlyProAsnSerPheGlnProSerAsp 132
321 GATGGANACTTATGGGAAGAGGTCTTAACAGATTTCACCCCTTCGGATCC 370
132 oMetLeuThMetAlaValAspCysGlyLeuLysAspAspLeuValLeuG 149
371 TAGTTGACATGCGTGTAGATTGTGGATTGAAGATGATTTGGTGTGG 420
149 LyAspProAsnAlaProArGpPheValLeuTrpLysGlyLysLeuArGp 165
421 GAGATCCGTATGGCGCTCTTGTCTTGTGGAAAGATTAACATAAGCCCT 470
166 ValProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePr 182
471 GTTCCCGGCAAGCTCACTGATCTCCCTTCTTGTGATGATGATGCC 520
182 oGlyLysLeuArGAlaGlyPheGlyProIleGlyLeuArGpProSerPr 199
521 TGGCAAGCTCAGAGCTGGTTTGTGCCCATTTGGCCCTTCGCCCTTGCCTC 570
199 roGlyHsGluGluSerValGlu 206
571 CAGGTTATGAGATCATCATTGAG 593
seq_name: gb_est50:AM695211

```

```

seq_documentation_block:
LOCUS AM695211 640 bp mRNA EST 15-JUN-2000
DEFINITION NP092606ST1P1051 Developing stem Medicago truncatula cDNA clone
ACCESSION AM695211
VERSION AM695211.1 GI:7569973
KEYWORDS EST.
SOURCE Medicago truncatula
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
1 (bases 1 to 640)
He,X.-T., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
,C.J., Flores,H.R., Imman,J.T., Weller,J.W., May,G.D. and Dixon
,R.A.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
TITLE Medicago truncatula stem library
JOURNAL Unpublished (2000)
COMMENT Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 640 Std Error: 0.00
Plate: 092 row: G column: 06
Seq primer: TCACACGAAACGCTATGAC.
FEATURES
Location/Qualifiers
source
1..640

```

```

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NP092606ST"
/tissue_type="stem"
/dev_stage="Pooled developmental"
/Vector="Lambda Zap; Contains a mixture of
internodal stem segments"
BASE COUNT 174 a 126 c 146 g 194 t
ORIGIN
alignment_scores:
Quality: 838.00 Length: 214
Ratio: 4.169 Gaps: 1
Percent Similarity: 93.925 Percent Identity: 76.636
alignment_block:
US-09-508-418-2 x AM695211 ..
Align seq 1/1 to: AM695211 from: 1 to: 640
300 AlaIleSerAlaArGleuGlySerLysLeuLysLeuSerTrpLysLeu 316
1 GCATTTTCCTGCTAGCTAGGTAACAAGGTAAGTATCTTGGAGCTTTT 50
316 rSerIleThrLysSerGluLysGlyGlyTyrlHisLeuThrTrpGluTrp 333
51 AAGTATTAGTAAACGTGAGATGGAGATCAGTTCGTTGAGTAAACCC 100
333 roGluGlyValAlaSerLeuGlnSerArGSerIleValMetThrValPro 349
101 CAGAAAGAGTGTCTTTCGACAGCAAGACTGTTATCATGCATTCCT 150
350 SerTyrlValAlaSerAsnIleLeuArGpProLeuSerValAlaAlaAlaAs 366
151 TCCCAAGTTGCTAGTCCATTTGCTGCACCTGCTCTGCTGCAGCA 200
366 PalaLeuSerAsnPheTyrlProProValGlyAlaValThrIleSerT 383
201 TGCACCTTCAAGTTTATTTATTCCTCCAGCTTCGCCAGTTCATTTCT 250
383 yrProGlnGluAlaIleArGAspGluArGLeuValAspGlyGluLeuLys 399
251 ATCCAAAGAGCCATTTAGATGCAATGTGTGATGATGAGGAGTTGAA 300
400 GlyPheGlyGlnLeuHisProArGTrpGlnGlyValGluTrpLeuGlyTh 416
301 GGATTTGGTCAATTTGCATCCAGCCAGGAGGATGCAACAACATTAAG 350
416 rIleTyrlSerSerLeuPheProAsnArGAlaProLysGlyArGValI 433
351 TATATACACCTCATCCCTTTTCCCTAACCCAGCACCACTCGAAGGGTTC 400
433 eudeuLeuAsnTyrlIleGlyAlaLysAsnProGluIleLeuSerLys 449
401 TACTCTTGAATTAACATTTGAGGGCTTACCACATTCCTGGATTTTAAAG 450
450 ThrGluSerGlnLeuValGluValAlaLysArGAspLeuArGlyLysMetI 466
451 ACGGAGGTGGAACCTTGTGAAGCAGTGTGAGATTTGGAAGAAAACCT 500
466 uIleTyrlProLysAlaGlnAspProLeuValValGlyValArGValTrpP 483
501 TATAAACCAGAAATGCTCAGATCCCAATTTTGGGGGTTAGAGTGGCC 550
483 roGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThr 499
551 CTCAGACTTATCCACAGTTCCTTATGAGCAATCTTGTGATGATGAGT 600
500 AlaLysAlaAlaMetAsnAspAsnGlyLeuGlnGlyLeu 512
601 GCTAAGTTCCTCTAATAAATACTGGAITTTGAGGGGCTT 640

```

seq_name: gb_est199:BG320666

seq_documentation_block:

LOCUS BG320666 780 bp mRNA EST 27-FEB-2001
DEFINITION Zm04_06e01.R.zm04.AAPC.ECORC_cold_stressed_maize_seedlings Zea mays
CDNA_clone Zm04_06e01, mRNA sequence.

ACCESSION BG320666 GI:13150344

VERSION BG320666.1

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.

REFERENCE Singh, J.A., Makui, K., Couroux, P., De Moors, A., Harris, L.J., Hattori
1 (bases 1 to 780)

AUTHORS Singh, J.A., Makui, K., Couroux, P., De Moors, A., Harris, L.J., Hattori

TITLE Expressed Sequence Tags from Cold-Stressed Maize Seedlings

JOURNAL Unpublished (2001)

COMMENT Contact: Singh, J.A.
Eastern Cereal and Oilseed Research Centre
960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.

FEATURES Location/Qualifiers

source 1..780

/organism="Zea mays"

/db_xref="taxon:4577"

/clone="Zm04_06e01"

/note="Vector: Bluescript SK-/XhoI-EORI; Site_1: Eco RI;
Site_2: Xho I; Lower temperature 50 C / hour from 22 to
120C; bring to 50 in 1 hour from 120C. Leave at 50C 2 days
, photoperiod 16 hours. Light intensity was 125 ue-1.
Library prepared by in vivo mass excision from amplified
library."

BASE COUNT 185 a 198 c 186 g 175 t 36 others

ORIGIN

alignment_scores:
Quality: 822.50 Length: 258
Ratio: 3.880 Gaps: 2
Percent Similarity: 82.171 Percent Identity: 67.054

Alignment_block:
US-09-508-418-2 x BG320666 ..

Align seg 1/1 to: BG320666 from: 1 to: 780
123 Glycerol-3-phosphate 5-phosphatase [Zea mays]
1 GGTCCCAACAGGCTCCAGCCGCCCTTCACCCATGCGCGTTGA 50

139 PCysglyLeuLysAspPaspLeuValLeuGlyAspProAsnAlaProArgp 156
1 GGTCCCAACAGGCTCCAGCCGCCCTTCACCCATGCGCGTTGA 100

156 heValLeuTrpLysGlyLysLeuArgProValProSerLysLeuThrAsp 172
101 TCGTCCTCTKGGAGGGGAAAGCGYGARCGCCGTCWCCAAACCCGCCGCGT 150

173 LeupropheasapleuMetSerLleProGlyLysLeuArgAlaGlyPh 189
151 CTCGGTTCCTCGAATCATGATGATCCAGGAGGAGCCAGCGCGGKCR 200
189 eglProIleGlyLeuArgProSerProGlyHisGluGluSerValG 205

201 AGCGCGCYTTGCAATCCGCCGCTCCYCCAGAGCGCCCAAGATYARPTG 250

206 IuGlnPheValAlaArgAsnLeuGlyGluValPheGluArgLeuIle 222

251 AGAAGTTCGTTGCCCAACCCYTTGGTGGAGKCTTGAGGSCCTCAAT 300

223 GluProPheCysSerGlyValIyrValGlyAspProSerLysLeuSerMe 239

301 KAGCCTTCTGCTCARGTGTCTATTCGTGGTGTCTTAAKCYCAGCAT 350

239 tLysAlaAlaPheGlyLysValItrPlyLysLeuGluGluThrGlyGlySeri 256

351 GAAGGCTCATTTGGGAAAGGTTTKCGGTTGGAAANAACCTGAGGTACTA 400

256 IleIleGlyGlyThrPheLysAlaIleLysGluArgSerSerThrProLys 272

401 TTATTKGTGAMC.ATCAAGACAAATTCARAGAGAGAGAGAGAGAGAGAA 449

273 AlaProArgAspProArgLeuProLysProLysGlyGlnThrValGlySeri 289

450 CCACCGADGAGATGCGCTTCCGAGGCAAAAGGCAAGACAGATTCGCATC 499

289 PheArgLysGlyLeuArgMetLeuProAspAlaIleSerAlaArgLeu 306

500 TTTGAGGAAGGGTCTGGCCAMGCTTCCAAATTCAAARACATCCAGCTTGG 549

306 IySerLysLeuLysLeuSerTrpLysLeuSerSerIleThrLysSerGlu 322

550 GTAGTAAAGTCAAACTATCAGGAAACCTCACGACGATTAACAAATCAGAT 599

323 LysGlyGlyThrHisLeuThrThrGlyGluThrProGluGlyValAlaSerIle 339

600 GACAAAGGATATGTTGGAGTATTAAGCCAAAGGGGTGTTGTTGGT 649

339 uGlnSerArgSerIleValMetThrValProSerTrpValAlaSerAsnI 356

650 GCAGGCTAAAGATTTTAMCATGACTATTCATATGTTGCTAGACACA 699

356 leIeuArgProLeuSerValAlaAlaAlaAspAlaLeuSerAsnPhery 372

700 WTTTCGCTCCACTTCAAGCATGCTGCAGAGCCCTATC...AAGATCTA 746

372 TrpProProValGlyAlaVal 379

747 TTAACCAACCGGTTGCTCCGGA 768

seq_name: gb_est199:BG352959

seq_documentation_block:

LOCUS BG352959 590 bp mRNA EST 01-MAR-2001
DEFINITION sab93f10.y1 gm-cl040 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: gm-cl040-2588 5, similar to SW:PROC_ARATH P55826
PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.

ACCESSION BG352959 GI:13181619

VERSION BG352959.1

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE Shoemaker, R., Kelm, P., Vodkin, L., Espelding, J., Coryell, Y., Khanna
1 (bases 1 to 590)

AUTHORS Shoemaker, R., Kelm, P., Vodkin, L., Espelding, J., Coryell, Y., Khanna

Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Perrson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project
Unpublished (1999)

TITLE Shoemaker R/Public Soybean EST Project

JOURNAL Public Soybean EST Project

COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134
Call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com High quality sequence stop: 481.

FEATURES
source
1_590
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1040-2588"
/lab_host="DH10B"
/issue_type="Hypocotyl and Plumule, germinating seeds"
/notes="Vector: pT733pac (Pharmacia); site:1: EcoRI; site_2: NotI; this cDNA library was constructed from mRNA isolated from hypocotyl and plumule tissues of seeds germinated for three days of the cultivar Williams 82. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestions with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the pT733-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

BASE COUNT 156 a 117 c 147 g 170 t
ORIGIN
Alignment_scores:
Quality: 797.00 Length: 191
Ratio: 4.478 Gaps: 1
Percent Similarity: 93.194 Percent Identity: 80.105

alignment_block:
US-09-508-418-2 x BG352959
Align seg 1/1 to: BG352959 From: 1 to: 590
190 GYpPOLLGLYLeuAArgProSerProProGlyHisGluGluSerValGI 206
115 GGTGGCCTTGGAAATTCGGCCCTCCTCCAGGTCAMGAGGAATCGGTGA 64
206 uGlnPheValArgArgAsnLeuGlyGlyValPheGluArgLeuIleG 223
223 IuProPheCysSerGlyValItyrValGlyAspProSerIysLeuSerMet 239
115 AGCCTTTTGTTCAGGGGCTCTAATGACGAGATCCCTCAAAAATTAAGTATG 164
240 LysAlaIaIaPheGlyIysValTrrpLysLeuGluGluThrGlyGlySerI 256
165 AAAGCGCAATTCGGAAAGTTTGGAAAGCTGGAATAAAATVHGGRGATC 214
256 eilleglylgytthPheLysAlaIaIeLysGluArgSerSerThrProLysA 273
215 TATTGGTGGAACTTTCAAGCAATACAAAGAGCAATATGGAAGCTTCAAAAC 264
273 IapRoArgAspProArgLeuProLysProLysGlyGlnThrValGlySer 289
265 CACTCTGGAGATCCGGCTGCGCAAAAACCAAAAGTCAAGCTGTGGATCT 314
290 PheArgLysGlyLeuArgMetLeuProAspAlaIleSerAlaIaArgLeuGI 306
315 TTCGGAAAGGACTTACACATGTTGCTGATGCAATTTTGCAGACTAGG 364

306 ySerLysLeuLysLeuSerTriplyLysLeuSerIlePrrLysSerGIul 323
365 CAACAAAGTAAGATTATCTTGGAGAGCTTCCAAAGTATNTGTAACAGCGAA 414
323 yScLysGlyItyrHisIstLeuThrTyrGluThrProGluGlyValValSerLeu 339
415 GTGAGAGATNACAGTTTGACATATGAAACACCGAAAGAGAGCTGTTCTTGG 464
340 GlnSerArgSerIleValMetThrValProSerTyrValAlaSerAsnII 356
465 CAGTCGAAAACGTGTGTCCGACCAATTCCTTCTATGTTGCTTAGTACAT 514
356 eleuArgProLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheryrT 373
515 GTCGGCGTCTGTGTCGTGGCTGTGCTGCAAGTCAACTTCAAGTATTAT 564
373 yrPro.ProValGlyAlaVal 379
565 ACCCTTCACAGTGTGCTGCAAGT 585

seq_name: gb_est74:BE455801
seq_documentation_block: 580 bp mRNA EST 27-FEB-2001
LOCUS BE455801
DEFINITION HVSMEG0015D05f Hordeum vulgare pre-anthesis spike EST library
HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone
HVSMEG0015D05f, mRNA sequence.
ACCESSION BE455801
VERSION BE455801.2 GI:13154672
KEYWORDS EST.
SOURCE
ORGANISM Hordeum vulgare
barley.
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 580)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Fritsch,D., Yu
,Y., Anderson,H., Dale,J., Henry,D., Kennodie,S., Palmer,M., Rambo
,T., Saski,C., Schwartzbeck,J., Simons,J., Choi,D.W., Main,D. and
Wood,T.
Development of a genetically and physically anchored EST resource
for barley genomics
Unpublished (2000)
On Jul 26, 2000 this sequence version replaced gi:9465836.
JOURNAL
COMMENT Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTAACCTCACAATAAGGG
High quality sequence stop: 499.
FEATURES
source
1_580
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="HVSMEG0015D05f"
/lab_host="SOLR"
/issue_type="pre-anthesis spike"
/notes="Vector: lambdaZAP; site_1: EcoRI; site_2: XhoI"
BASE COUNT 156 a 124 c 150 g 147 t
ORIGIN
alignment_scores:
Quality: 778.00 Length: 192
Ratio: 4.346 Gaps: 0
Percent Similarity: 93.229 Percent Identity: 75.000

alignment block:

US-09-508-418-2 x BE455801 ...
Align seg 1/1 to: BE455801 from: 1 to: 580

217 ValPheGluArgLeuIleGluProPheGlySerGlyValIleTyrValGlyAs 233
3 GTCCTTTGAGCGCTATGCAACCTTCTGCTCAGGCTGATGCTGGTGGTGA 52
233 pProSerIysLeuSerMetIysAlaAlaPheGlyLysValIleTyrPheLys 250
53 TCCTTCCAAAGCCCAATATGAAAGGCTGCTTTGGAGAGGTTGGAGGTTGG 102
250 IuGIuThrGlyGlySerIleIleGlyGlyIlePheLysAlaIleLysGlu 266
103 AGGAATTTGGGGGTAGTATTTGGTGGAAACCAACCAAGGCCATTCAGAGT 152
267 ArgSerSerThrProLysAlaProArgAspProArgLeuProLysProly 283
153 AAAGGAAAGAAACCCCAACCCGCAAGGATCCCGACTTCGGCAGCAACAA 202
283 sGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspA 300
203 GGGACAGACGGGTGCATCTTTCAGAAAGGCTCTGGCCATGCTCCGAAATG 252
300 lAlIeSerAlaArgLeuGlySerIysLeuLysLeuSerTyrPheLysLeuSer 316
253 CCATGCGCATGAGTGGGTAGTAAAGTCAAGACGCTCATGGAAACCTTACG 302
317 SerIleThrIysSerGlyLysGlyIleTyrHisLeuThrTyrGluThrPr 333
303 AGCATTTCAAAAGGGGCAACCAAGGATATGCTATAGGTTATGAAACACC 352
333 oGluGlyValValSerLeuGlnSerArgSerIleValMetIleThrValPro 350
353 AGAAGGACTGTTCAGTGCAGCCTAAAGGTTATCATCATGACCATCCCGT 402
350 eTyrValAlaSerAsnIleLeuArgProLeuSerValAlaIleAlaAsp 366
403 CATTTGTGCTAGTATCTTACGCCACTTTCATTAATGATGCGAGCAT 452
367 AlAlaSerAsnPheTyrTyrProProValAlaGlyAlaValThrIleSerTyr 383
453 GCACCTGCAAAATTCATTTATCCGCCAGTTCGCTTAACCTGTTTATATA 502
383 rProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysG 400
503 TCCAAANAGAGCTATTTAGAAAGAAATGCTTNAATGATGGGAAAGCTCAGG 552
400 lYpHeGlyGlnLeuHisProArgThr 408
553 GGTTCGGCCAGCTGATCCACGATAGC 578
seq_name: gb_est39:AV538014

seq_documentation block:
LOCUS AV538014 596 bp mRNA EST 06-SEP-2000
DEFINITION AV538014 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION CDNA clone RZ109a12F 3', mRNA sequence.
VERSION AV538014
KEYWORDS AV538014.1 GI:8698297
SOURCE EST.
ORGANISM Arabidopsis thaliana
thale cress.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 596)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)

MEDLINE 20363093
CONTACT: Erika Asamizu
COMMENT The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
LOCATION/Qualifiers
SOURCE 1..596
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="RZ109a12F"
/clone_lib="Arabidopsis thaliana roots Columbia"
/issue_type="roots"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 165 a 135 c 126 g 170 t
ORIGIN

alignment_scores:
Quality: 739.00 Length: 177
Ratio: 4.452 Gaps: 0
Percent Similarity: 93.785 Percent Identity: 79.096

alignment_block:
US-09-508-418-2 x AV538014/reverse

Align seg 1/1 to reverse of: AV538014 from: 1 to: 596
372 TyrTyrProProValGlyAlaValThrIleSerTyrProGlnGluAlaI 388
591 TATTTACCCACACAGTTCACAGCATATCTATCTGATCCGCAAGAAAGCAAT 542
388 eArgAspGluArgLeuValAspGlyGluLeuLysGlyPheGlyGlnLeuH 405
541 CGGAACAAGAAATGTTGATGATGAGTGAACTAAAGGTTTTGGGCAATTTGC 492
405 lSProArgThrGlnGlyValGluThrLeuGlyThrIleTyrSerSerSer 421
491 ATCCACGACAGCGAAGGATTTGAAACATTTAGGAACATATCTACAGCTCCCA 442
422 LeuPheProAsnArgAlaProLysGlyArgValLeuLeuLeuAsnTyrI 438
441 CTCCTTCCAAATCCGCGCACCGCCGGAAAGAAATTTGCTGTGAACCTACAT 392
438 eGlyValAlaLysAsnProGluIleLeuSerIysThrGlnSerGlnLeuY 455
391 TGGCGGCTCTACAAACACCGGAATTCCTGTCCAAGTCTGAAAGGTGAGTTAG 342
455 aGlnuValValAspArgAspLeuArgLysMetLeuIleYsProLysAla 471
341 TGGAAAGCAAGTTGACAGAGATTTGAGGAAATGCTTAATTAAGCTTAATTCG 292
472 GlnAspProLeuValValGlyValArgValTyrProGlnAlaIleProG 488
291 ACCGATCCACTTAATTTAGAGCTTAGGGTATGCGCTCAAGCCATTCCTCA 242
488 nPheLeuValGlyHisLeuAspThrLeuSerThrIleLysAlaIleAlaMet 505
241 GTTCTAGTGGTGCATTTGATATCTTGTACAGCGGCTTAATTAATTCGTA 192
505 snAspAsnGlyLeuGlnGlyLeuPheLeuGlyLysTyrValSerGly 521
191 GCTCTTCCGGCTAACGAAGGCTAATTTTGGGGGCAATTTAGCTGCTCGTGG 142
522 ValAlaLeuGlyArgCysValGluGlyValArgValTyrGluValAlaSerGlu 538
141 GTAAGCTTAGGGGCTGATGAGAAAGCCATATTTGAAACCGGATGAGGT 92
538 lThrGlyPheLeuSerArgTyrAlaTyrIys 548
91 CAACAACCTTCAATGTCACGGTACGCTTACAG 61

```

seq_name: gb_est190:BF624024
seq_documentation_block:
LOCUS       BF624024         510 bp    mRNA            EST          21-FEB-2001
DEFINITION   HVSMba0008116f Hordeum vulgare seedling shoot EST library
             HVCNDNA0001 (Cold stress) Hordeum vulgare cDNA clone HVSMba0008116f,
             mRNA sequence.
ACCESSION   BF624024
VERSION     BF624024.2     GI:13082749
KEYWORDS    EST.
SOURCE      barley.
             Hordeum vulgare
REFERENCE   Ekarlyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
             ; Triticeae; Hordeum.
             1 (bases 1 to 510)
AUTHORS     Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Fritsch, D., Yu
             , Y., Anderson, H., Dale, J., Henry, D., Kerrodle, S., Palmer, M., Rambo
             , T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and
             Wood, T.
TITLE       Development of a genetically and physically anchored EST resource
             for barley genomics
             Unpublished (2000)
JOURNAL     On Dec 18, 2000 this sequence version replaced gi:11887758.
COMMENT    Contact: Wing RA
             Clemson University Genomics Institute
             Clemson University
             100 Jordan Hall, Clemson, SC 29634, USA
             Tel: 864 656 7288
             Fax: 864 656 4293
             Email: rwing@clemson.edu
             Seq primer: AATTAACTTCACTCAATAAGGC
             High quality sequence stop: 505.
             Location/Qualifiers
             1..510
             /organism="Hordeum vulgare"
             /cultivar="Morex"
             /db_xref="taxon:4513"
             /clone="HVSMba0008116f"
             /HVCNDNA0001 (Cold stress)"
             /tissue-type="Seedling shoot"
             /lab_host="TJC121"
             /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT  148 a      110 c      116 g      135 t      1 others
ORIGIN
Alignment_scores:
Quality:     685.00      Length:     168
Ratio:       4.363      Gaps:       0
Percent Simlarity: 93.452      Percent Identity: 77.381
alignment_block:
US-09-508-418-2 x BF624024 ..
Align seg 1/1 to: BF624024 from: 1 to: 510
320 LysSerGluLysGlyGlyTyrHisIleuThrTyrGluThrProGluGlyVala 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 AAGCGCGCAACCAAGGATATGTATTAGGTATGAAACACCCGGAAGGACT 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
336 lValSerLeuGlnSerArgSerIleValMetThrValProSerTyrVala 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 TGTTCAGAGCGAGGCTAAAGGTTATCATGACCAATCCCGTCATATGTTG 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
353 lAsenAnlleuAtrgProLeuSerValAlaAlaAlaAspAlaLeuSer 369
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 CRAAGATATCTTACGCCCACTTTCATTTGATGACGAGATGCACACTCA 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 AspPheTyrTyrProProValAlaValAlaThrIleSerTyrProGlnGI 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

156 AAATTCATTAATCCCGCAGTGTGCTGTAACGTATTCAATCCAAAGA 205
386 uAlAlIeaTgaspGluArgLeuValAspGlyIuleuLysGlyPheGlyG 403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 AGCTATTAGAAAAGAATGCTTAATGATGGGAGCTCCAGGGTTCGGCC 255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
403 lInleuHisProArgThrGlnGlyValaIgluThrleuGlyThrIleYSer 419
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 AGCTGCATCCACCGTACCCAGAGATGCGAGACTTTAGGCAATATATAC 305
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 SerSerLeuPheProAsnArgAlaProLysGlyArgValleuLeuLeu 436
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 TCTTCTCTCTTTCCTCCCAATCGTCTGCTCGGAAAGGTTTCTCTGTA 355
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 nYrYlIeGlyGlyAlaLysAsnProGluIleLeuSerLysThrGluSerG 453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 CTTATATCGGGGGTTCACAAATATACAGGATCGTCCAAAGACGAGATG 405
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
453 lInleuValGluValAlaAspArgAspLeuArgLysMetLeuIleLysPro 469
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
406 ACTTATAGAGAAAGCTGTGATCGTGCATCTCAGAAATAATGTTGAATACC 455
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 LysAlaGlnAspProLeuValValGlyValArgValItrProGlnAlaI 486
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
456 AGAGCAGCAGACCCCTTATAGCATTAGGGGTCAGAGTGTGGCCACCAACAT 505
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
486 ePro 487
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
506 ACCA 509
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: gb_est42:AM119581
seq_documentation_block:
LOCUS       AM119581         498 bp    mRNA            EST          17-JUL-2000
DEFINITION   sdd4d01.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
             Gm-c1016-3002 5' similar to SW:PROCARRATN P53826 PROTOPORPHIRINOGEN
             OXIDASE, CHLOROPLAST PRECURSOR ; mRNA sequence.
ACCESSION   AM119581
VERSION     AM119581.1     GI:6094914
KEYWORDS    EST.
SOURCE      soybean.
ORGANISM   Glycine max
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
             1 (bases 1 to 498)
AUTHORS     Shoemaker, R., Keim, P., Vodkin, L., Erpelund, J., Coryell, V., Khanna
             , A., Bolla, B., Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
             Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
             , Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuirk
             , R., Riltter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
             , R., Materston, R. and Wilson, R.
             Public Soybean EST Project
TITLE       Unpublished (1999)
JOURNAL     Public Soybean R/Public Soybean EST Project
COMMENT    Contact: Shoemaker R/Public Soybean EST Project
             Public Soybean EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@wustl.edu
             This clone is available through: Genome Systems, Inc. 4633 World
             Parkway Circle St. Louis, Missouri 63134 For further information
             call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
             427-3324 or contact: clones@genomesystems.com or
             info@genomesystems.com web site: www.genomesystems.com
             Insert Length: 1114      Std Error: 0.00
             Seq primer: -40RP from Gibco
             High quality sequence stop: 425.
             Location/Qualifiers
             1..498
             /organism="Glycine max"
             /db_xref="taxon:3847"
             source

```

```

/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-3002"
/clone_lib="Gm-c1016"
/tissue_type="Immature flowers of field grown plants"
/lab_host="X110-Gold"
/notes="Vector: pBluescript II XR, Site_1: EcoRI, Site_2: XhoI. This cDNA library was constructed from mRNA isolated from Immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into X110-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."
BASE COUNT      137 a      106 c      111 g      143 t      1 others
ORIGIN

```

```

alignment_scores:
  Quality: 674.00      Length: 164
  Ratio: 4.266        Gaps: 0
  Percent Similarity: 96.341      Percent Identity: 82.317
alignment_block:
  US-09-508-418-2 x AW119581 ..
  Align seg 1/1 to: AW119581 from: 1 to: 498

```

```

328 LeuTTRTYRGIUHHPrroGluGlyValValSerLeuGlnSerArgSerI1 344
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 107
58  TGCTCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 107
361 erValAlaAlaAlaAspaLalaLeuSerAsnphetyrTYRPrroValAlGly 377
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 157
108 CTGCGTGTGCTGCAGATGCACCTTCAAAAGTTTATTTACCCCGCAGTTGCT 157
378 AlaValAlaTrpIleSerTYRPrroGlnGluAlaIleArgAspGluArgLeuVa 394
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 207
158 GCAGTTTCATATCTTATCCCAAAAAGAGCTATTAGATCAGCAAAATGCTTGAT 207
394 IAspDlYgluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGly 411
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 257
208 AGATGGTGTAGTTGAAGGGGTTTGGTCAATTCGATCCACGTRAGCCAAAGAG 257
411 aGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAla 427
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 307
258 TGGAAACAATAGGAACATATATACAGCTCATCATTTCCCAACCCGAGCA 307
428 ProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnPr 444
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 357
308 CCACCTGGAAAGGGTTTCACTCTTGAATTTACATTTGGAGGAGCAACTAATAC 357
444 OGIlleLeuSerLysTrpGlnSerGlnLeuValGluValAlaAspArg 461
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 406
358 TGGAAATTTTATCGAAGCGACAGCATGTGAACr. GTGGAAACAGTTGATGAG 406
461 sPLeuATrGlySMetLeuIleLysProLysAlaGlnAspProLeuValVal 477
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 456
407 ATTTGAGAGAAAATTCCTTATTAATAAATCCCAATGCCCCAGATTTGTAGTN 456
478 GLyValArgValAlaTrpProGlnAlaIleProGlnPheLeuVal 491
      |||TTCCTGACCAATTCCTTCTTATGTTGGTAGACATGTGGCTCCTCGT 498
457 GGGGTGAGACTGTGGCTCAAGCTAATTCACACAGTTTCTTAAAGTT 498
seq_name: gb_est69:BE021686

```

```

seq_documentation_block:
  LOCUS      BE021686      480 bp      mRNA      21-NOV-2000
  DEFINITION sm61b05.y1 Gm-c1028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
  Gm-c1028-8362 5' similar to SW:PPCCARRAH P55826 PROTOPORPHYRINOGEN
  OXIDASE, CHLOROPLAST PRECURSOR ; mRNA sequence.
  ACCESSION BE021686
  VERSION   BE021686.1 GI:8284127
  KEYWORDS EST.
  SOURCE     soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
            1 (bases 1 to 480)
  REFERENCE Shoemaker R., Kelm, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna
            A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
            Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
            Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
            R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
            R., Materston, R. and Wilson, R.
            Public Soybean EST Project
            Unpublished (1999)
  JOURNAL   Contact: Shoemaker R./Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
  COMMENT   This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 1299 Std Error: 0.00
            High quality sequence stop: 394.
            Location/Qualifiers
            1..480
            /organism="Glycine max"
            /db_xref="taxon:3847"
            /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-8362"
            /clone_lib="Gm-c1028"
            /tissue_type="roots of 'supernod' plants"
            /lab_host="DH10B"
            /note="Vector: pBluescript II XR, Site_1: EcoRI, Site_2:
            XhoI. The mRNA was isolated from roots of Glycine max
            'Supernod' plants generously donated by Dr. Gary Stacey.
            The seedlings were inoculated with Bradyrhizobium
            japonicus, strain USDA110 prior to harvest. Stratagene's
            cDNA synthesis kit (catalog number 200401) was used to
            synthesize the cDNA. First-strand synthesis was performed
            with 5-methyl dCTP, hence the ligated cDNA was
            hemimethylated. A modification of Stratagene's
            first-strand synthesis primer was used. An 'anchor'
            nucleotide (V=A,C, or G) was added to the 3' end of the
            primer [GAGAGAGAGAGAGAGAGAGACTAGTCGTGAG(T)18V] to anchor
            the primer at the 5' end of the poly(A) tract. After
            second-strand synthesis, the cDNA ends were filled in with
            cloned Pfu DNA polymerase, ligated to EcoRI adapters and
            subsequently phosphorylated. The XhoI site within the
            first-strand synthesis primer was then restricted by
            digestion with XhoI; all XhoI sites in the cDNA would be
            protected by their hemimethylated status. The cDNA
            constructs were size-fractionated with a 500bp cutoff,
            using GibcoBRL Life Technologies' cDNA Size Fractionation
            column. The column eluent was then ligated into
            Stratagene's pBluescript II XR predigested vector
            (pBluescript II SK(+)) that has been digested with EcoRI
            and XhoI, and phosphorylated by Stratagene). Both the
            white and blue colonies appear to contain recombinant
            plasmids with cDNA inserts, based on size (n=25). This
            library was constructed by Dr. Paul Kelm and Dr. Virginia

```

BASE COUNT 138 a 98 c 104 g 137 t 3 others

alignment_scores: Quality: 629.00 Length: 158 Ratio: 4.338 Gaps: 0 Percent Similarity: 91.772 Percent Identity: 77.848

alignment_block: US-09-508-418-2 x BE021686 .. Align seg 1/1 to: BE021686 from: 1 to: 480

297 LeuproaspalaialeSerataarJeuGlySerlyseuylseuSerrtr 313
7 TTGGCTGATGCAATTTTCCAGACTAGCCAAACAAGTAATATCTG 56
313 plysleuSerSerIleThrLysSerGluuLygLYTYRHIsleuThrT 330
57 GAAGGTTTCAAGTATTAAGTAACTGATAGTGGAGAGTACGTTTGACAT 106
330 YRGLuThrProGluuLygLYValValSerLeuGlnSerArgSerIleValMet 346
107 ATGNAACACCGAAGAGAGTGGTTTTCAGTGCAGCAAAACTGTCTCCTG 156
347 ThrValProSerTyrrValValSerAsnIleLeuArgProLeuSerValAl 363
157 ACCATTCCTTCTATGTTGCTAGTACATGCTGCTCTGCTGCTGCTG 206
363 aAlaAlaaspAlaLeuSerAsnPhetrrTyrrProProValGlyValValT 380
207 TGCTGCAAGTGGACTTGAAGTTTATTAATACCTCCAGTTGCTGCAAGTTT 256
380 hrIleSerTyrrProGlnIuAlaIleArGAspGluArGLeuValAspGly 396
257 CCATATTCCTATCCAAAGAGAGCTATTAGATCAAAATGCTGTATGATGAT 306
397 GluLeuLygLYPheGlyGlnLeuHisProArgThrGlnGlyValGluThr 413
307 GAGTTGAAAGGGGTTGGTCAATTTGCATCCAGTACCCAAAGAGGGAAGC 356
413 rLeuGlyThrIleTyrrSerSerSerLeuPheProAsnArgAlaProLySg 430
357 ATTAGAACATATATACAGCTCATCTATTCCTCCCAACCGAGACACCACTG 406
430 lYArgValleuLeuLeuAsnTYRllegLYGlyAlaLysAsnProGluIle 446
407 AAAGGTTTCTACTCTNGAATTTACATTTGGAGGAGCAACTAATAATCTGGAAT 456
447 LeuSerLyThrGluSerGlnLeu 454
457 NATATCAAAAGCGACAGTGAACCTT 480
seq_name: gb_est39:BG355389

seq_documentation_block: LOCUS BG355389 479 bp mRNA EST 06-MAR-2001 DEFINITION 947043H05.y2 947 - 2 week shoot from Barkan lab Zea mays CDNA, mRNA

ACCESSION BG355389 KEYWORDS BG355389.1 GI:13237375 SOURCE EST

ORGANISM Zea mays. Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea. 1 (bases 1 to 479) REFERENCE 1 Malbot, V. Maize ESTs from various cDNA libraries sequenced at Stanford University

JOURNAL: Unpublished (1999) COMMENT: Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 947043 row: H column: 05.

FEATURES source 1..479 /organism="Zea mays" /cultivar="B73" /db_xref="taxon:4577" /clone_lib="947 - 2 week shoot from Barkan lab" /rissue_type="leaf and stem, including leaf base" /dev_stage="2 week Old seedling (3 leaves)" /lab_host="XLI-Blue" /note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-); Site.1: EcoRI; Site.2: XhoI; Directionally cloned using Stratagene's Unizap XR cDNA cloning kit with the 5' end at the EcoRI site. The library represents 8 x 10e5 independent recombinant phage. The plants were greenhouse grown."

BASE COUNT 122 a 108 c 126 g 123 t ORIGIN

alignment_scores: Quality: 621.00 Length: 151 Ratio: 4.500 Gaps: 0 Percent Similarity: 91.391 Percent Identity: 79.470

alignment_block: US-09-508-418-2 x BG355389 .. Align seg 1/1 to: BG355389 from: 1 to: 479

398 LeuLygLYPheGlyGlnLeuHisProArgThrGlnGlyValGluThrle 414
2 CTCACGAGGCTTTGGCCAGTGGCATCCACCGATGCAAGGAGTTGAGACAT 51
414 uGlyThrIleTyrrSerSerSerLeuPheProAsnArgAlaProLySgLYA 431
52 AGGAACAAATATACAGTTCCTACTCTTCCAAATTCCTGCTCCGACGGTA 101
431 rYValleuLeuLeuAsnTYRllegLYGlyAlaLysAsnProGluIleLeu 447
102 GGGGTACTCTTAACACTCATAGAGGAGTCTCAAAACACAGCAATTTGTT 151
448 SerLyThrGlnLeuSerGlnLeuValGluValAlaLysPArgAspLeuArgLY 464
152 TCCAGAGCTGAAAGTGAAGTGGTGCATATCAGTTGACCCGTTGACCTCCG 201
464 sMetLeuIleLysPProLySAlaGlnAspProLeuValValGlyValArgY 481
202 AATGCTTATTAATTTCTACAGCAGTGCACCTTATGATTCCTGGTTCGAG 251
481 altTpproGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeu 497
252 TTGGCCACAAGCCATATCTCAAGTCTCTGTTAGGACATCTTGATCTTTG 301
498 SerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLe 514
302 GAAGCCCGCAAAAGCTGCCAGCCGAGGCTGACCTGAGCTGTTCCCT 351
514 uGlyGlyAsnTYrValSerGlyValAlaLeuGlyArgGlyValGluGlyVA 531
352 AGGAGGGAACATATGTTGCAGAGTGGCTGCGGAGATGGTGGTGAAGGCG 401
531 lArYrGlyValAlaSerGlyValThrGlyPheLeuSerArgTYrAlaTYR 547
402 CGTATGAAGTGGCTCCGAATATCTGACTTCTTGGACCAAGTATMGCTTAC 451

548 Lys 548
452 AAG 454

seq_name: gb_est39:AV540305

seq_documentation_block:
LOCUS AV540305 519 bp mRNA EST 06-SEP-2000
DEFINITION AV540305 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION AV540305
VERSION AV540305.1 GI:8702063
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 519)

REFERENCE
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093
COMMENT Contact: Erika Asamizu
The First Laboratory For Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
Location/Qualifiers
1..519
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="R2148e10F"
/clone_lib="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 150 a 116 c 109 g 144 t
ORIGIN
alignment_scores:
Quality: 605.00 Length: 146
Ratio: 4.416 Gaps: 0
Percent Similarity: 93.836 Percent Identity: 78.767

alignement_block:
US-09-508-418-2 x AV540305/rev ..
Align seg 1/1 to reverse of: AV540305 from: 1 to: 519

403 GlnLeuHisProArgThrGlnGlyValAlaGluThrLeuGlyThrIleLeuPhe 419
519 CAATTGGATCGCGGAGCAAGGATTGAAACATTTGGAACTATCTACAG 470
419 rSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuLeuLeu 436
469 CTCCTCACTCTTTCCAAATCGCGACCGCCGGAAGAAATTTGCTGTGA 420
436 snfYrIleGlyValAlaLysAsnProGluIleLeuSerLysThrGluSer 452
419 ACTACATTTGGGGGTCTACAAACCGCAATTCCTGCCAAGTCTGAAGGT 370
453 GlnLeuValAlaValAlaValAspArgAspLeuArgLysMetLeuIleLysP 469
369 GAGTGTAGTGAAGAGAGTGGACAGAGATTGAGGAATAATGTAATTAAGCC 320
469 GysAlaIleGlnAspProLeuValAlaGlyValArgValTTPProGlnAlaI 486

319 TAATTCGACCGATCCACTTAATTAAGAGTTAGGGTATGGCCCTCAAGCCA 270
486 leProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysAla 502
269 TTCCTCACTCTTTCCAAATCGCGACCGCCGGAAGAAATTTGCTGTGA 220
503 AlameLAsnAspAsnGlyLeuGlyLeuPheLeuGlyLysAsnYrVa 519
219 TCCTCAAGCTCTTCGGGCTACGAAGGCTATTTTGGGTGGCAATTTACGT 170
519 lSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAla 536
169 CGCTGTGTAGCCCTTACGCGCGGTGTGAAGAGCGCCATATGAAGCCGCA 120
536 ergIValThrGlyPheLeuSerArgTyrAlaTyrLys 548
119 TTGAGGTCAACACCTTATGTACAGGTACCGCTTACAG 82

seq_name: gb_est24:A1748531

seq_documentation_block:
LOCUS A1748531 457 bp mRNA EST 17-JUL-2000
DEFINITION sb54d01.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl016-314.5 similar to SW:PPCC_ARATH_P55826 PROTOPORPHRINOMGEN
OXIDASE, CHLOROPLAST PRECURSOR ; mRNA sequence.
ACCESSION A1748531
VERSION A1748531.1 GI:5126795
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 457)

REFERENCE
AUTHORS Shoemaker R., Keim, P., Vodkin, L., Erpeliding, J., Coryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
TITLE Shoemaker R/ Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 1096 Std Error: 0.00
High quality sequence stop: 401.
Location/Qualifiers
1..457
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-314"
/clone_lib="Gm-cl016"
/tissue_type="immature flowers of field grown plants"
/lab_host="X110-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally

FEATURES
source

cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into X110-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Expelling.

BASE COUNT 124 a 96 c 106 g 131 t

alignment_scores: Quality: 593.00 Length: 152
Ratio: 4.090 Gaps: 1
Percent Similarity: 95.395 Percent Identity: 80.263

alignment_block: US-09-508-418-2 x A1748531 ..
Align seg 1/1 to: A1748531 from: 1 to: 457

335 G1YVALVALSERLEUGINSEARVGSERILEVALMETHYVALPROSERTY 351
3 GGAGTGGTTCTTTCAGTCCAAACACCTGCTCCGACCAATCCCTCCTTA 52
351 YVALALSERASNILEUAVRQPROLEUSERVALALALASPAAL 368
53 TGGTGTAGACATTTGCTGCTCCTCCTCCTGCTGCTGCTGACATGAC 102
368 euSerAsnPhetYrTYrPrOvAlGlyAlaValThrIleSerTYrPro 384
103 TTTCAAAGTTTAAATACCCTCCAGTGTGCTGCATATTCATTCCTATCCA 152
385 G1NG1UALALALVARGASPLUARGLARGLVAVLALASP 6LYGLUENLUSG1YP 401
153 AAGGAAAGCTAATTAAGATGCAAGAAATGCTGATGAGATGGTGAAGGGCT 202
401 heGlyGlnLeuHisprOArGrInglyValGluThrLeuGlyThrIle 417
203 TTTGCATATTCATCCACACGTCAGCCAGAGAGTGAAGAACAATA 252
418 TyrSerSerLeuPheProAsnArgAlaProlYsGlyAraGValLeuLe 434
253 TACACCTCATATCATAATCCCAACCGACACACCCTCGAAGAGGTTCTACT 302
434 uLeuAsnTYrIleGlyGlyAlaValAsnProGluIleLeuSerIysPheG 451
303 CTTGAATTACATTTGAGAGCACTAACTAGGAATTTTATCCGAAGAGG 352
451 IuSerGlnLeuValGluValAlaAsparGAspLeuArgLysMetLeuIle 467
353 ACAGTGAACCT 6TGGAAACAGTTGATCGAGATTTTGAGGAAATAATCCTTATA 401
468 LysProLysAlaGlnAspProLeuValValGlyValAraGValTTPrpG 484
402 AACCCCAATGCCACAGATTCATTTTGTAGTGGGGTGAAGAGCGCTCA 451
484 nAlA 485
452 AGCT 455

seq_name: gb_est197:BG133906

seq_documentation_block: 440 bp mRNA EST 31-JAN-2001
LOCUS BG133906 EST466798 tomato crown gall Lycopersicon esculentum cDNA clone
DEFINITION EST466798 tomato crown gall Lycopersicon esculentum cDNA clone

ACCESSION BG133906
VERSION BG133906.1 GI:12634094
KEYWORDS EST.
SOURCE tomato
ORGANISM Lycopersicon esculentum
Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

REFERENCE 1 (bases 1 to 440)
AUTHORS van der Hoeven, R., Sun, H., Cho, J., Uteerback, T., Hansen, C., Romling, C., and Tanksley, S.
TITLE Generation of ESTs from tomato crown gall tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
Location/Qualifiers
1..440
/organism="Lycopersicon esculentum"

FEATURES
source
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cT0E14B22"
/clone_lib="tomato crown gall"
/tissue_type="crown gall"
/dev_stage="crown galls from full-grown plants (8 wks old)"
/lab_host="SOLAR"
/note="Vector: bluescript SK(-). Site_1: EcoRI; Site_2: XhoI; Four wk old greenhouse plants were stab inoculated on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr, Cornell U.). Galls were allowed to develop for another 4 wks, when gall tissue was frozen in liquid nitrogen."

BASE COUNT -104 a 105 c 109 g 122 t

alignment_scores: Quality: 589.00 Length: 145
Ratio: 4.675 Gaps: 2
Percent Similarity: 86.897 Percent Identity: 77.931
alignment_block: US-09-508-418-2 x BG133906 ..
Align seg 1/1 to: BG133906 from: 1 to: 440
1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGln.. 16
5 ATGACAAACAACAGCCGCTGCATCAACCATCTTACGATTTTCCACCCGGTGC 54
17serserserserPro...LeuAlaP 24
55 GCCGCTGCCCTGCGCCGCCCTCCCTCCATCCATCCGCTATTTTAT 104
24 heLeuAsnArGrThrSerPheIleProPheSerSerIleSerIysArGAsn 40
105 TTTTAAATCGTACGAAATTTATTCATCCTTTCCACCCTCCACCCGCAAGC 154
41 SerValAsnGlyAsnGlyTPrPArGrThrArgGlySerValAlaLysAspTy 57
155 AGTGTCATATTCGATATGCGTGGAGAAACCGGTGTCGTTGCGAAGAATTA 204
57 rThrValProSerSerAlaValAspGlyGlyProAlaAlaGluLeuAspC 74
205 TACAGTTTCCCTCCTCAGAACTGACGGTAACTGATACCCGGAGCTGGATT 254
74 ysValIleValGlyAlaGlyIleSerGlyLeuGlyIleAlaGlnValMet 90
255 GTGGGTAGTTCGGAGCGAAGTAATTAAGTGTCTCCATTTGCTAAGTGATA 304
91 SerAlaAsnTYrProAsnLeuMetValThrGluAlaArGAspArgAlaG 107
305 TCGGCTAATTAATCCCAATTTGATGGTGAAGCGAGGCGGAGATGCTCCGG 354
107 yGlyAsnIleThrThrValGluLuarGAspGlyTYrLeuTrpGluGluGlyP 124
355 TGCAACAATATACGACGGTGAAGAGATGATGATCTTAATGGGAAGAAGGTC 404

124 r0AsnSerPheGlnProSerAspPheMetLeuThr 135
 |||
 405 CTAACAGTTTCCAGCCCTTCGGATCCTATGTTGACT 439
 seq_name: gb_est73:BE326089

seq_documentation_block:
 LOCUS BE326089 590 bp mRNA EST 21-DEC-2000
 DEFINITION NF084G10ST1F1081 Developing stem Medicago truncatula cDNA clone
 VERSION NF084G10ST 5', mRNA sequence.
 BE326089
 BE326089.2 GI:11936674
 EST:
 SOURCE barlet medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
 1 (bases 1 to 590)
 He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
 ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
 ,R.A.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Expressed sequence tags from the Samuel Roberts Noble Foundation
 Medicago truncatula stem library
 Unpublished (2000)
 On Jul 14, 2000 this sequence version replaced gi:9199866.
 Contact: Dixon RA
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7302
 Fax: 580 221 7380
 Email: radixon@noble.org
 Medicago Genome Initiative accession: MGI:S:26673
 Insert Length: 703 Std Error: 0.00
 Plate: 084 row: 6 column: 10
 Seq primer: TCACACGAGAAACACGCTATGAC.
 Location/Qualifiers
 1..590
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone_lib="NF084G10ST"
 /clone_lib="Developing stem"
 /tissue_type="stem"
 /dev_stage="Pooled developmental"
 /note="Vector: Lambda zap; Contains a mixture of
 internodal stem segments"
 BASE COUNT 130 a 153 c 131 g 176 t
 ORIGIN

alignment_scores: Quality: 552.50 Length: 198
 Ratio: 3.635 Gaps: 7
 Percent Similarity: 76.768 Percent Identity: 59.596
 alignment_block:
 US-09-508-418-2 x BE326089 ..
 Align seg 1/1 to: BE326089 from: 1 to: 590

8 AsnHisProAsnIlePheThrHisGlnSerSerSerSerProLeuAlaIap 24
 :|||
 21 TCTCCACCCCT...CTTTCGACCCACGCTAACCCCTTCTCCGCCCTTA..... 62
 24 eLeuAsnArgThrSer.PheIleProPheSerSer..... 35
 63TCTCCATTCCTCCCTTCCCTCCCAATTCGAATAATCA 99
 36IleSerIysArgAsnSerValAsnArgSAsnGlyTrpArgThrAr 50
 100 AAATTCATTCGCAATCGCAATTCACCATTAAC...CCCATTCCTCCGCTCCG 146

50 gCysSerValAlaLysAspTyrThrVal.....ProSerSera 63
 |||
 147 TTGCTCATTCGGGAGAAATCCACCCGATCTCCGACTAAACACCTCA 196
 63 lAvAlAspGlyGlyProAlaAlaGluLeuAspCysValIleValAla 79
 :|||
 197 AATCAAAATCCGGAGAGAGTCTCCTCATCCGTTGATTTGTGGTGGT 246
 80 GlyIleSerGlyLeuSerIleAlaGlnValMetSerAlaAsnTyr.... 94
 |||
 247 GGCATTCAGTGCCCTTGGCATCCGCTCAGGCTCTTCTACAAAGCATCTCA 296
 95ProAsnLeuMetValYhrGluAlaArGAsPArGAlaGly 109
 |||
 297 TGGTGTCCCAATGTTGTTTTCACAGAGGCAAGCCAGTTGGTGGGA 346
 109 snIleThrThrValGluArGAsPglyTyrLeuTrpGluGlyProAsn 125
 |||
 347 ATATTATTACTGTGGAGAGAGATGGGTTTCTTGGGAAAGAGTCCATA 396
 126 SerPheGlnProSerAspPheMetLeuThrMetAlaValAspCysGly 142
 |||
 397 AGTTTTCAMCCTTCTGATGCTATGCTACCACATGTTGGTGAACATG 446
 142 uLysAspAspLeuValLeuGlyAspProAsnAlaProArGpPheValLeu 159
 |||
 447 GAAGGACGAMCTTGTGTTGGTGCATCCCTGATGCACCTAATTTGGTGT 496
 159 rPlYsGlyLysLeuArGProValProSerLysLeuThrAspLeuProPhe 175
 |||
 497 GGAATGGGAAATGAGCCGCGGCAAAACCAACTGATTTGGCTTTC 546
 176 PheAspLeuMetSerIleProGlyLysLeuArgAlaGlyPhe 189
 |||
 547 TTTGACTTAAATGATTTTGTGGCAAGATCAAGAGCTGGCTTT 588
 seq_name: gb_est51:AW755783

seq_documentation_block:
 LOCUS AW755783 451 bp mRNA EST 21-NOV-2000
 DEFINITION s109gc1.y1 Gm-cl036 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl036-813 5' similar to SW:PPCC_ARAHH_P53826 PROTOPORPHYRINOGEN
 OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
 ACCESSION AW755783
 VERSION AW755783.1 GI:7685135
 EST:
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
 1 (bases 1 to 451)
 REFERENCES
 1 Shoemaker,R., Keim,P., Vodkin,L., Erpedding,J., Coryell,V., Khana
 ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers
 ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Halvey,N., Schurk
 ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 ,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134 For further information
 call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or
 info@genomesystems.com web site: www.genomesystems.com
 Insert Length: 894 Std Error: 0.00

FEATURES High quality sequence stop: 298.
 Location/Qualifiers
 source 1..451
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1036-813"
 /clone_lib="Gm-c1036"
 /issue_type="somatic embryos cultured on MSD 20"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lilia Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

BASE COUNT 125 a 88 c 111 g 127 t
 ORIGIN

alignment_scores: Quality: 543.00 Length: 146
 Ratio: 4.022 Gaps: 0
 Percent Similarity: 92.466 Percent Identity: 75.342

alignment_block:
 US-09-508-418-2 x AW755783 ..

Align seg 1/1 to: AW755783 from: 1 to: 451

```

382 SerTYrPrroGInGluAlAlleArGspGluArgLeuValAspGlyGluLe 398
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
11 TCCTGTCCAAAGAGACTATATAGATCAGAAATGCTTGTATGATGGTGA 60
398 ULySGlYPhEGlYGlNLeuHISPrOArGTYrGInGlyAlGluThrLeuG 415
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
61 GAAGGGGTTTGGTCAATTGCATCCAGTACGCAAGGAGACTGGAAACAT 110
415 lYThrlleTySerSerleuPheProAsnArgAlaProLYSGlYArg 431
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
111 GAACrATATACAGCTCATCCATrATTCCCAACCGAGCACCACTGGAA 160
432 ValLeuLeuAsnTYrIleGlyValAlALySAsnProGluIleLeuSe 448
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
161 GTTGTACTCTGTAAATTAATGTGGAGAGACTAACTGCAATTTTATTC 210
448 rLySThrGluSerGlnLeuValGluValAlAspArgAspLeuArgLYSM 465
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
211 AAAGAGGCACTGAACCTGTGGAACAAGCTGATCGAGATTGGAGAA 260
465 eLleuIleLYSPrOlySAlAGlnAspProLeuValValGlyValArgVal 481
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
261 TCCTrATTAACCCAAATGCCCCAGATCCATTGTAGTGGGGGTGAGACT 310
482 TrPrroGInAlAllePrroGInPheLeuValGlyHISLeuAspThrLeu 498
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
311 TGGCCrTAAGCTTATTCACAGCTCTTATTTGGCCATATTGATCTCTGA 360
498 rThrAlAlALySAlAlAlMetAsnAspAsnGlyLeuGluIleuPheLeuG 515
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
361 TGTGGCTAAAGCTTCTATCAAGAAATACGTGGTTTGAAGGGCTCTTTCT 409
515 lYGlYAsnTYrValSerGlyValAlAlALeUglYArgCYs 527
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
410 G.GGGACTTATGTGTCTGGCGGCTGTACTTATGACGATGC 447

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:33:02 ; Search time 100.76 Seconds
(without alignments)
719.563 Million cell updates/sec

Title: US-09-508-418-2
Perfect score: 2844
Sequence: 1 MTTPTIAPHNPITFTHOSSSS.....EGAVYVASEWYTGILSRVAYK 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- 1: sp.archaea:**
- 2: sp.bacteria:**
- 3: sp.fungi:**
- 4: sp.human:**
- 5: sp.invertebrate:**
- 6: sp.mammal:**
- 7: sp.mhc:**
- 8: sp.organelle:**
- 9: sp.phage:**
- 10: sp.plant:**
- 11: sp.podent:**
- 12: sp.unclassified:**
- 13: sp.vertebrate:**
- 14: sp.virus:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2832	99.6	548	10	O9SBI4
2	2824	99.3	548	10	O9SAX9
3	2565.5	90.2	557	10	O64384
4	2061	72.5	545	10	O9ZS15
5	2004.5	70.5	555	10	O9SPL6
6	1989.5	70.0	535	10	O9M629
7	1966.5	69.1	562	10	O9LRI8
8	1530.5	53.8	563	10	O9ZTAY
9	1527.5	53.7	563	10	O9P0P9
10	522.5	18.4	502	10	O9S1W5
11	515	18.1	436	2	O67814
12	459	16.1	544	10	O9P1V8
13	456	16.0	501	10	O9LVA7
14	456	16.0	547	10	O9FMS9
15	424	14.9	404	10	O64385
16	407.5	14.3	467	2	O9KDX8
17	363.5	12.8	462	2	O9KRV9
18	317	11.1	468	2	O69829
19	290.5	10.2	391	10	O9FWS3

RESULT ID	PRELIMINARY:	PRT:	548 AA.	ALIGNMENTS	
20	254	8-9	435	2	O9J540
21	250	8.8	475	5	O9UBD0
22	249	8.8	435	5	O92718
23	247	8.7	475	5	O9VC52
24	236.5	8.3	63	10	O9M5P0
25	231.5	8.1	424	2	O9PLH9
26	211	7.4	424	2	O84750
27	174.5	6.1	421	2	O9RSV7
28	169.5	6.0	441	2	O9RPN3
29	162	5.7	524	13	O918A7
30	153.5	5.4	416	2	O9R154
31	148	5.2	440	1	O58162
32	146	5.1	430	1	O9HRU3
33	144	5.1	488	10	O9LVT1
34	143	5.0	490	10	O9SKX5
35	136.5	4.8	467	2	O9S2P5
36	130	4.6	890	2	O9VW97
37	127	4.5	574	10	O9FVR9
38	126	4.4	489	2	O9KHE1
39	126	4.4	499	2	O9K566
40	125	4.4	578	10	O9FV57
41	121.5	4.3	443	2	O82865
42	121.5	4.3	479	2	O9R6X5
43	121	4.3	404	1	O28536
44	119.5	4.2	472	2	O9ZBG4
45	119.5	4.2	570	10	O65813

RESULT ID	SEQUENCE	548 AA.	59241 MW:	00E1A106F299E4AF	CRC64:
AC	O9SBI4				
DT	01-MAY-2000 (TRMBLrel. 13, Created)				
DT	01-MAY-2000 (TRMBLrel. 13, Last sequence update)				
DE	PROPOPHYRINOLIN OXIDASE PX-1 (EC 1.3.3.4).				
OS	Nicotiana tabacum (Common tobacco).				
OC	Eukaryota: Viridiplantae: Embryophyta: Tracheophyta; Spermatophyta:				
OC	Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;				
OC	Solanales; Solanaceae; Nicotiana.				
OX	NCHI_TaxID=4097;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. SRI;				
RA	Horikoshi M., Mametsuka K., Hirooka T.;				
RT	"The Molecular Basis of Photobleaching Herbicide Resistance in Tobacco."				
RT	Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.				
DR	EMBL: AF044128; AAD02290.1; -				
KW	Oxidoreductase.				
SQ	SEQUENCE 548 AA: 59241 MW: 00E1A106F299E4AF CRC64;				

Query Match	Best Local Similarity	99.6%	Score 2832;	DB 10;	Length 548;
Matches 546;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
OY	1	MTPTIAPHNPITFTHOSSSPPLAFINRTSFTIPFSSISKRNSVNCNGMFRKCSVAKDYTP 60			
DB	1	MTPTIAPHNPITFTHOSSSPPLAFINRTSFTIPFSSISKRNSVNCNGMFRKCSVAKDYTP 60			
OY	61	SSAVYGGPAALDVCYIVAGISGLCIAQVMSANYPNLMVTEARBARAGSNTTVRRDGLW 120			
DB	61	SSAVYGGPAALDVCYIVAGISGLCIAQVMSANYPNLMVTEARBARAGSNTTVRRDGLW 120			
OY	121	EEGPNFQPPSDPMLTMAVDCGLKDDLVLDGPNAPRFVLMKSKLRVPSKLLDLPFFDLMS 180			
DB	121	EEGPNFQPPSDPMLTMAVDCGLKDDLVLDGPNAPRFVLMKSKLRVPSKLLDLPFFDLMS 180			
OY	181	IPGKLRAGFGPIGLRSPSPGHEBSEYEQFVRKNLGGVFERLIPEFCGQVYVGDPSKLSMK 240			

```

181 |IPKLRAGFGAIGLRPSPGHESEVQFVRRNLGGVEFERLIEPFCSGGVYVAGDPSKLSMK 240
241 |AAGKWKLEETGGSIIGTFEKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
241 |AAGKWKLEETGGSIIGTFEKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
301 |ISARLGSKLLKLSWKLSSTTKSEKGYHLTYETPEGVVSLQSRSTVMTPVPSVAANILRPL 360
301 |ISARLGSKLLKLSWKLSSTTKSEKGYHLTYETPEGVVSLQSRSTVMTPVPSVAANILRPL 360
361 |SVAADALSNFYPPVGAVTISYPOEAI RDERLVDGELKGGQLHPRTOGVEITLGTIYSS 420
361 |SVAADALSNFYPPVGAVTISYPOEAI RDERLVDGELKGGQLHPRTOGVEITLGTIYSS 420
421 |SLFPNRAPKRGVLLNLYIGGAKNPEILSKTESQLEVVDRDLRKMILKPKAQDPLVYGV 480
421 |SLFPNRAPKRGVLLNLYIGGAKNPEILSKTESQLEVVDRDLRKMILKPKAQDPLVYGV 480
481 |VMPQAIPOFLVGHLDLTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRVCVGAIVEVASEVTG 540
481 |VMPQAIPOFLVGHLDLTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRVCVGAIVEVASEVTG 540
541 |FLSRYAYK 548
541 |FLSRYAYK 548

```

RESULT 2

```

09SAX9 PRELIMINARY; PRT: 548 AA.
ID 09SAX9
AC 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2001 (TREMBlrel. 16, Last annotation update)
DE PLASTIDAL PROTOPORPHYRINOGEN OXIDASE.
GN NMPPOX1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAISON NR;
RX PubMed=9608719;
RA Matanabe N., Che F., Iwano M., Takayama S., Nakano T., Yoshida S.,
RA Isogai A.;
RT "Molecular characterization of photomixotrophic cultured tobacco cells
RT resistant to protoporphyrinogen oxidase-inhibiting herbicides.";
RL Plant Physiol. 118:751-756(1998).
DR EMBL: AB020501; BAA34713.1; -.
DR InterPro: IPR000894; -.
DR Prodom: PD000290; -.
SQ SEQUENCE 548 AA: 59215 MW: AB51A513FD3CETBC CRC64;

```

```

Query Match 99.3%; Score 2824; DB 10; Length 548;
Best Local Similarity 99.3%; Pred. No. 7.6e-200;
Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1 MTTTPIAHNPITFHSSSPLAFLNRTSFTPFSSISKRNSVNCNGMRTKRSVAKDVTVP 60
1 MTTTPIAHNPITFHSSSPLAFLNRTSFTPFSSISKRNSVNCNGMRTKRSVAKDVTVP 60
1 MTTTPIAHNPITFHSSSPLAFLNRTSFTPFSSISKRNSVNCNGMRTKRSVAKDVTVP 60
61 SSAVDGGPAEIDCVYAGAGISGLCIAQVMSANYPNLMVTEARDRAGNITTYVERDGYLW 120
61 SSAVDGGPAEIDCVYAGAGISGLCIAQVMSANYPNLMVTEARDRAGNITTYVERDGYLW 120
61 SSAVDGGPAEIDCVYAGAGISGLCIAQVMSANYPNLMVTEARDRAGNITTYVERDGYLW 120
61 SSAVDGGPAEIDCVYAGAGISGLCIAQVMSANYPNLMVTEARDRAGNITTYVERDGYLW 120
121 BEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKLTLDFPFDLMS 180
121 BEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKLTLDFPFDLMS 180
121 BEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKLTLDFPFDLMS 180

```

```

181 |IPKLRAGFGAIGLRPSPGHESEVQFVRRNLGGVEFERLIEPFCSGGVYVAGDPSKLSMK 240
181 |IPKLRAGFGAIGLRPSPGHESEVQFVRRNLGGVEFERLIEPFCSGGVYVAGDPSKLSMK 240
241 |AAGKWKLEETGGSIIGTFEKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
241 |AAGKWKLEETGGSIIGTFEKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
301 |ISARLGSKLLKLSWKLSSTTKSEKGYHLTYETPEGVVSLQSRSTVMTPVPSVAANILRPL 360
301 |ISARLGSKLLKLSWKLSSTTKSEKGYHLTYETPEGVVSLQSRSTVMTPVPSVAANILRPL 360
301 |ISARLGSKLLKLSWKLSSTTKSEKGYHLTYETPEGVVSLQSRSTVMTPVPSVAANILRPL 360
361 |SVAADALSNFYPPVGAVTISYPOEAI RDERLVDGELKGGQLHPRTOGVEITLGTIYSS 420
361 |SVAADALSNFYPPVGAVTISYPOEAI RDERLVDGELKGGQLHPRTOGVEITLGTIYSS 420
421 |SLFPNRAPKRGVLLNLYIGGAKNPEILSKTESQLEVVDRDLRKMILKPKAQDPLVYGV 480
421 |SLFPNRAPKRGVLLNLYIGGAKNPEILSKTESQLEVVDRDLRKMILKPKAQDPLVYGV 480
481 |VMPQAIPOFLVGHLDLTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRVCVGAIVEVASEVTG 540
481 |VMPQAIPOFLVGHLDLTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRVCVGAIVEVASEVTG 540
541 |FLSRYAYK 548
541 |FLSRYAYK 548

```

RESULT 3

```

064384 PRELIMINARY; PRT: 557 AA.
ID 064384
AC 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BINTJE;
RA Johnston D.J., Droz E., Rochaix J.D., Malnoe P.C.;
RL Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) -> PROTOPORPHYRIN-
CC IX + H(2)O.
DR EMBL: AJ225107; CAA12400.1; -.
DR Mendel; 29307; Soltu:3076;29307.
DR Oridoreductase.
KW Oridoreductase.
SQ SEQUENCE 557 AA: 60463 MW: AE2B15B17E0B89F8 CRC64;

```

```

Query Match 90.2%; Score 2565.5; DB 10; Length 557;
Best Local Similarity 88.9%; Pred. No. 8.7e-181;
Matches 495; Conservative 21; Mismatches 32; Indels 9; Gaps 2;
1 MTTTPIAHNPITFHQ-----SSSSP-LAFLNRTSFTPFSSISKRNSVNCNGMRTKRC 51
1 MTTTPIAHNPITFHQ-----SSSSP-LAFLNRTSFTPFSSISKRNSVNCNGMRTKRC 51
1 MTTTPIAHNPITFHQ-----SSSSP-LAFLNRTSFTPFSSISKRNSVNCNGMRTKRC 51
52 SVAKDVTVPSSAVDGGPAEIDCVYAGAGISGLCIAQVMSANYPNLMVTEARDRAGNIT 111
52 SVAKDVTVPSSAVDGGPAEIDCVYAGAGISGLCIAQVMSANYPNLMVTEARDRAGNIT 111
52 SVAKDVTVPSSAVDGGPAEIDCVYAGAGISGLCIAQVMSANYPNLMVTEARDRAGNIT 111
52 SVAKDVTVPSSAVDGGPAEIDCVYAGAGISGLCIAQVMSANYPNLMVTEARDRAGNIT 111
112 TYBRDGYLWEEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKL 171
112 TYBRDGYLWEEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKL 171
112 TYBRDGYLWEEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKL 171
112 TYBRDGYLWEEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKL 171
121 TYBRDGYLWEEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKL 180
121 TYBRDGYLWEEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKL 180
121 TYBRDGYLWEEGPNSSFPSPMILTMVAVDGLKDDLVLDGPNAPRFVLMKGLRPPVSKL 180
172 DLFPFDLMSIPKLRAGFGAIGLRPSPGHESEVQFVRRNLGGVEFERLIEPFCSGGVYV 231
172 DLFPFDLMSIPKLRAGFGAIGLRPSPGHESEVQFVRRNLGGVEFERLIEPFCSGGVYV 231

```

Db 181 DLPEFLMSTIPKLRAGFGAIGLRPSPGYESVQFVVRNMGAEVFERLLIEPFCSCGYA 240

QY 232 GDSKLSMKAARFKVYKLEBETGSSIIIGTFKAIKERSSTPKAPRDPRLPKPGQGVGSR 291

Db 241 GDSKLSMKAARFKVYKLEBETGSSIIIGTFKAIKERSSTPKAPRDPRLPKPGQGVGSR 300

QY 292 KGIKRLMPLDIAISARLGSKLSMKSSTTSEKGGYHLTYETPEGVVSLQSRIVMTVPSY 351

Db 301 KGIKRLMPLDIAISARLGSKLSMKSSTTSEKGGYHLTYETPEGVVSLQSRIVMTVPSY 360

QY 352 VASNIIRPLSVAAADALSNPPYPPVGAVTISYPOEAIIRBERLVDGELKFGQLHRTQGV 411

Db 361 VASNIIRPLSVAAADALSNPPYPPVGAVTISYPOEAIIRBERLVDGELKFGQLHRTQGV 420

QY 412 ETLGTYSSSLFPNRPAPKRVLLNNTYGGAKPPELITKTESOLVYEVVDRLKMLIKPKA 471

Db 421 ETLGTYSSSLFPNRPAPKRVLLNNTYGGAKPPELITKTESOLVYEVVDRLKMLIKPKA 480

QY 472 QDBPLVGVVWPQALIPQFLVGHLDLSTAKAAMNDNGLEGLFLGNYVSGVALGRVESA 531

Db 481 QDBPLVGVVWPQALIPQFLVGHLDLSTAKAAMNDNGLEGLFLGNYVSGVALGRVESA 540

QY 532 YEVASEVTGFLSRAYK 548

Db 541 YEVASEVTGFLSRAYK 557

RESULT 4

Q9ZS15 PRELIMINARY; PRF: 545 AA.

AC Q9ZS15; 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)

DE T15B16.13 PROTEIN.

GN T15B16.13.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eumstoids II;

OC Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA WASHU;

RT "the A. thaliana Genome Sequencing Project. ";

RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Stoneking T., Smith R.;

RT "the sequence of A. thaliana T15B16. ";

RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Waterston R.;

RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF104919; AAC72870.1; -

DR Mendei: 39261; AArch; 3076; 39261.

DR InterPro: IPR000209; -

DR InterPro: IPR001613; -

DR PRINTS: PRO0757; AMINEOXDASEF.

DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.

SO SEQUENCE 545 AA; 58598 MW; 6AB825691A01BE70 CRC64;

Query Match 72.5%; Score 2061; DB 10; Length 545;

Best Local Similarity 74.3%; Pred. No. 1.3e-143;

Matches 404; Conservative 49; Mismatches 73; Indels 18; Gaps 5;

QY 22 LALNRTSFLPFSISKRNSVNCNGM---RTRCSVAKDYTVVSSAVDGGPAEEL--DCVI 76

Db 24 SGGSLTSKRNPRYLITYSPAHRK---CNRRWRFCSIAKSDSPITPISNFSQPLDCVI 79

Db 3 LSLRPTTQSLPSPSKRN-LRLAVYKPLRLRCVYAGGPTVYSSKISGGGTTITDCCI 61

QY 77 VGAGISGLCIAQVMSANY-----PMLMVTIARDRAGNTITVERDYLMEBGPNSFPSPD 132

Db 62 YGGGISGLCIAQALATAKHPDAPNLIYTEAKDRYGNITTEENGFMEBGPNSFPQSDP 121

QY 133 MFLMAYDGLKDLVLDGPNAPRRLVLMKGLRVPYSKLTDLPFDLMSIPGKLRAGGPI 192

Db 122 MLTMVVDGLKDDLDVLDGPDTPARFVLMWNGKLRPVPSKLTLDLPFDLMSIGKIRAGFAL 181

QY 193 GLRPPRPHESVQFVVRNMGAEVFERLIEPFCSCG-----VYVGDSPKLSMKAARF 244

Db 182 GTRSPPRESESVFVVRNMGAEVFERLIEPFCSGRVIDKTSVYAGDSDSKSMKAARF 241

QY 245 KWKLEEFGGSSIIIGTFKAIKERSSTPKAPRDPRLPKPGQGVSPFKGLMPLDAISAR 304

Db 242 KWKLEEFGGSSIIIGTFKAIKERSSTPKAPRDPRLPKPGQGVSPFKGLMPLDAISAR 301

QY 305 LGSKIKLSMKSSTTSEKGGYHLTYETPEGVVSLQSRIVMTVPSYVANSIIRPLSVAA 364

Db 302 LGSKIKLSMKSSTTSEKGGYHLTYETPEGVVSLQSRIVMTVPSYVANSIIRPLSVAA 361

QY 365 ADALSNPPYPPVGAVTISYPOEAIIRBERLVDGELKFGQLHRTQGVETLGTYSSTLFP 424

Db 362 ANALSRLYPPVAAVSYSPKFAIRTECLIDGELKFGQLHRTQGVETLGTYSSTLFP 421

QY 425 NRAPKRVLLNNTYGGAKPPELITKTESOLVYEVVDRLKMLIKPKADPLVGVVWPQ 484

Db 422 NRAPKRVLLNNTYGGAKPPELITKTESOLVYEVVDRLKMLIKPKADPLVGVVWPQ 481

QY 485 AIPQFLVGHLDLSTAKAAMNDNGLEGLFLGNYVSGVALGRVESADEVYTGFLSR 544

Db 482 AIPQFLVGHLDLSTAKAAMNDNGLEGLFLGNYVSGVALGRVESADEVYTGFLSR 541

QY 545 YAYK 548

Db 542 YAYK 545

RESULT 5

Q9SP16 PRELIMINARY; PRF: 555 AA.

AC Q9SP16; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE PROTOPORPHYRINOGEN IX OXIDASE (EC 1.3.3.4).

GN PPIX.

OS Cichorium intybus (Chicory).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterales;

OC easterids II; Asterales; Asteraceae; Lactuceae; Cichorium.

OX NCBI_TaxID=13427;

RN [1]

RP SEQUENCE FROM N.A.

RA Adomat C., Boeger P.;

RT "Cloning, Sequence, Expression and Characterization of

RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF10961; AAF00194.1; -

DR InterPro: IPR000209; -

DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.

KM Oxidoreductase.

SO SEQUENCE 555 AA; 60244 MW; 0AB31FA9DDEB303D CRC64;

Query Match 70.5%; Score 2004.5; DB 10; Length 555;

Best Local Similarity 71.5%; Pred. No. 1.9e-139;

Matches 383; Conservative 60; Mismatches 84; Indels 9; Gaps 3;

QY 18 SSSPLAFLNRTSFLPFSISKRNSVNCNGMRTKCSVAKDYTV-SSAVDGGPAEELDCVI 76

Db 24 SGGSLTSKRNPRYLITYSPAHRK---CNRRWRFCSIAKSDSPITPISNFSQPLDCVI 79

QY 77 VGAGISGLCIAQVMSANY-----PNLMTVEARDRAGGNTTVERDGYLWEEGPNFSQSDP 132
 DB 80 VGAGISGLCIAQVMSANYSPDIVTVEARDRAGGNTTVERDGYLWEEGPNFSQSDA 139
 QY 133 MLTMAVDCGLKDDLVLDGPNAPRFVLMKGLRVPVSKLTDLPFPDLMSIPGKLRAGGPI 192
 DB 140 MLTMAVDCGLKDDLVLDGPNAPRFVLMKGLRVPVSKLTDLPFPDLMSIPGKLRAGGPI 199
 QY 193 GURPPEPHESVSEGFVAVRNLGGVEFERLIPFCSGVYVGGPASKLSMKAARAKWKLEBE 252
 DB 200 GURPPEPHESVSEGFVAVRNLGGVEFERLIPFCSGVYVGGPASKLSMKAARAKWKLEBE 259
 QY 253 GGSIIIGFTHAIAKERSSTPKAPRDLRPPKGGQVGFRRKGLRMLPDAISRLSKSLKLS 312
 DB 260 GGSIIIGFTHAIAKERSSTPKAPRDLRPPKGGQVGFRRKGLRMLPDAISRLSKSLKLS 319
 QY 313 WKLSTITSEKGGYHLYETEBGVVSLQSSRSITVMTPSYVANSNLRPLVSVAADALSIFY 372
 DB 320 WKLSTITSEKGGYHLYETEBGVVSLQSSRSITVMTPSYVANSNLRPLVSVAADALSIFY 379
 QY 373 YRPVAVVIISYPOEAIIRBERLVYDDELKGFGLHRPTGQVETLGTIYSSLPFNAPRGYV 432
 DB 380 YRPVAVVIISYPOEAIIRBERLVYDDELKGFGLHRPTGQVETLGTIYSSLPFNAPRGYV 439
 QY 433 LLNLTIGGAKNPEILSKTESQVLEVYVDRDLRKKMLIKPKAOPDLVYVGVWPFQAIPOFLV 492
 DB 440 LLNLTIGGAKNPEILSKTESQVLEVYVDRDLRKKMLIKPKAOPDLVYVGVWPFQAIPOFLV 499
 QY 493 HLDITLSTKAKAMNDNGLEGLFLGGNVYSGVALGRGVGAVYVASEVYGFSLRYAYK 548
 DB 500 HYDIISAKAKALSSGCFQGFGLGGNVYSGVALGRGVGAVYVASEVYGFSLRYAYK 555

RESULT 6
 Q9M629 PRELIMINARY; PRT; 535 AA.
 AC Q9M629;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PROTOPORPHYRINOGEN IX OXIDASE.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae;
 CC Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. B73 INBRED;
 RA Volrath S.L.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF218052; AAF26417.1; -
 DR SEQUENCE 535 AA; 56835 MW; 9EEFB0BE2973F919 CRC64;

Query Match 70.0%; Score 1989.5; DB 10; Length 535;
 Best Local Similarity 72.3%; Pred. No. 2,3e-138;
 Matches 394; Conservative 49; Mismatches 69; Indels 33; Gaps 6;

QY 17 SSSSEFLANRSTPSSSISKRNVCNGWRTRCSVAKDYTPVSSAVDGGPA----- 69
 DB 11 TAAAPL-LNSTR-IPALRLRHRGLSFR-----AAVAGGAABAAPASTG 50

QY 70 --AEIDYVAGVAGISGLCIAQVMSANY--PNLMTVEARDRAGGNTTVER--DGYLWEEG 123
 DB 51 ARLSADCVVAVGGISGLCTAQAALTRGVGVIVTEARARAGGNTTVEREEGYLWEEG 110

QY 124 PNSFQSDPMLTMAVDCGLKDDLVLDGPNAPRFVLMKGLRVPVSKLTDLPFPDLMSIPG 183
 DB 111 PNSFQSDPMLTMAVDCGLKDDLVLDGPNAPRFVLMKGLRVPVSKLTDLPFPDLMSIPG 170

QY 184 KLRAGFGPIGLRPSPGHESEVQFVRRNLGGEVFERLIEPFCSGVYVGGPASKLSMKAAR 243

DB 171 KLRAGFGALGIRPRPGRRESVEEAFVRRNLGAWEFERLIEPFCSGVYVGGPASKLSMKAAR 230
 QY 244 GKWKLEBEWGGSITIGTFEAIKERSSTPKAPRDLRPPKGGQVGFRRKGLRMLPDAISA 303
 DB 231 GKWVLEBEWGGSITIGTFEAIKERSSTPKAPRDLRPPKGGQVGFRRKGLRMLPDAISA 290
 QY 304 RLSKLSMKSLSITSEKGGYHLYETEBGVVSLQSSRSITVMTPSYVANSNLRPLVSVA 363
 DB 291 SLSKLSMKSLSITSEKGGYHLYETEBGVVSLQSSRSITVMTPSYVANSNLRPLVSVA 350
 QY 364 AADALSNEYFPYVAVTISYPOEAIIRBERLVYDDELKGFGLHRPTGQVETLGTIYSSLSLF 423
 DB 351 AADALSNEYFPYVAVTISYPOEAIIRBERLVYDDELKGFGLHRPTGQVETLGTIYSSLSLF 410
 QY 424 PNRAPRGYVILLNLTIGGAKNPEILSKTESQVLEVYVDRDLRKKMLIKPKAOPDLVYVGVW 483
 DB 411 PNRAPRGYVILLNLTIGGAKNPEILSKTESQVLEVYVDRDLRKKMLIKPKAOPDLVYVGVW 470
 QY 484 QAIPOFLVGHLDITLSTKAKAMNDNGLEGLFLGGNVYSGVALGRGVGAVYVASEVYGF 543
 DB 471 QAIPOFLVGHLDITLSTKAKAMNDNGLEGLFLGGNVYSGVALGRGVGAVYVASEVYGF 530

QY 544 RYAYK 548
 DB 531 RYAYK 535

RESULT 7
 Q9LR18 PRELIMINARY; PRT; 562 AA.
 AC Q9LR18;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (PROTOK-I).
 GN SO-POXI.
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
 CC Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. TONIC; TISSUE=LEAVES;
 RA Che F.S., Watanabe N., Iwano M., Inokuchi H., Takayama S., Yoshida S.,
 Isogai A.,
 RT "Molecular characterization and subcellular localization of
 RT protoporphyrinogen oxidase in spinach chloroplasts."
 RL Plant Physiol. 0:0-(2000).
 DR EMBL: AB029492; BAA96808.1; -
 DR InterPro: IPR001613; -
 DR InterPro: IPR001848; -
 DR PRINTS: PR00757; AMINEOXDASEF.
 DR ProDom; PD001272; -; 1.
 DR SEQUENCE 562 AA; 59929 MW; 638DF6E63D259C4 CRC64;

Query Match 69.1%; Score 1966.5; DB 10; Length 562;
 Best Local Similarity 71.1%; Pred. No. 1,2e-136;
 Matches 381; Conservative 58; Mismatches 84; Indels 13; Gaps 5;

QY 26 NR-TSPFPSSISK-R-NSVNCNGWRTRCSVA----KDYVSSAVD--GGPAALDVIY 76
 DB 27 NRITLILPSSSLRFRGSSIRCSSTISTSMSAANAANKNIGTNGVDGGGGGVLDVCI 86

QY 77 VGAGISGLCIAQVMSANYPNL--MTEARDRAGGNTTVERDGYLWEEGPNFSQSDP 132
 DB 87 VGGISGLCIAQVMSANYPNLSTNFTVTEAKDRVGGNTITMBAADGYLWEEGPNFSQSDA 146

QY 133 MLTMAVDCGLKDDLVLDGPNAPRFVLMKGLRVPVSKLTDLPFPDLMSIPGKLRAGGPI 192
 DB 147 VLTMAVDCGLKDELLVLDGPNAPRFVLMKGLRVPVSKLTDLPFPDLMSIPGKLRAGGPI 206

```

QY 193 GIRPSPGHESVDFQFVARNLNGEVEFRLLEPFCSSGYVYVGDPSKLSMKAAGKVMKLEET 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 207 GIRPSPGHESVDFQFVARNLNGEVEFRLLEPFCSSGYVYVGDPSKLSMKAAGKVMKLEEK 266
QY 253 GGSIIIGGFKAIKREKSSPPKAPRDPRLPKPKGQVWSPFRKGLRMLPRAISARLGSMTLKS 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 267 GGSIIIGGFKAIKREKSSPPKAPRDPRLPKPKGQVWSPFRKGLRMLPRAISARLGSMTLKS 326
QY 313 WKLSITSEKGGYHLYETPEPEGVVSLQSRISVMTVPSYVAASNILRPLSYAADAALSNFY 372
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 327 WLSGIKSSNGEYVNLTYETPDGLVYVTRKSVMTVPSYVAASSILRPLSDVAASELSLKFH 386
QY 373 YRPVAVNITVYPOAIRBERLVDELKFGQLHPRTOGVETLGTIYSSSLFPMRAKGRV 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 387 YRPVAVNITVYPOAIRBERLVDELKFGQLHPRTOGVETLGTIYSSSLFPMRAKGRV 446
QY 433 LLNLYIGGAKNPELISKTESOLVEVYVDRDLRKLKPKADPLVYVGRVWPAIQPFLV 492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 447 LLNLYIGGAKNPELISKTESOLVEVYVDRDLRKLKPKADPLVYVGRVWPAIQPFLV 506
QY 493 HLDPLSTAKAMNDNGLEGLFLGNYVSVVALGRVCGAVAVASEVTGFLSRVAYK 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 507 HFDLDAKAAKALTDGGRKGLFLGNYVSVVALGRVCGAVAVASEVTGFLSRVAYK 562

```

RESULT 8
Q9ZTA7 PRELIMINARY: PRT: 563 AA.

AC Q9ZTA7
DI 01-MAY-1999 (TRMBLrel. 10, Created)
DI 01-MAY-1999 (TRMBLrel. 10, Last sequence update)
DI 01-NOV-1999 (TRMBLrel. 12, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE PRECURSOR.
GN PPX1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaeae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC-621;
RX MEDLINE=99077317; PubMed=9862501;
RA Randalph-anderson B.L., Sato R., Johnson A.M., Harris E.H.,
RA Hauser C.R., Oeda K., Ishige F., Nishio S., Gillham N.W.,
RA Boynton J.E.;
RT "Isolation and characterization of a mutant protoporphyrinogen oxidase
RT gene from Chlamydomonas reinhardtii conferring resistance to porphyrin
RT herbicides."
RL Plant Mol. Biol. 38:839-858(1998).
DR EMBL: AF068635; AAC79685.1; -
DR Mendel: 39099; Chltre:3076;39099
SQ SEQUENCE 563 AA; 59802 MW; FC5E88F9C34CAFPC2 CRC64;

```

Query Match 53.8%; Score 1530.5; DB 10; Length 563;
Best Local Similarity 59.5%; Pred. No. 1.7e-104;
Matches 300; Conservative 73; Mismatches 114; Indels 17; Gaps 6;
QY 61 SSAVDGGPAE-----IDCVYVAGISGLCTAQVMSANY--PMLMTEARDRAG 107
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 60 AAAATGAPTAGAGAVAKLIDVWYVIVYVGGGLVYVGOALAAQHKIQNPLVTEARERVG 119
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 108 GNITTVERDGYLMEGRPSPDPMLTMAVDCGLKDDVLYLGPVNAPRFLYMKGKLRV 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 120 GNITVMSGDGYVMEGRPSPDPMLTMAVDCGLKDDVLYLGPVNAPRFLYMKGKLRV 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 168 SKLIDLPFFDLMSIPGKTRAGFGPGL- RSPSPGHEBSVDFQFVARNLNGEVEFRLLEPFC 226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 180 SGL-DALFFDLMSIPGKTRAGFGPGL- RSPSPGHEBSVDFQFVARNLNGEVEFRLLEPFC 238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 227 SGVYVGDPSKLSMKAAGKVMKLEETGGSIIIGGFKAIKREKSSPPKAPRDPRL- PKPKGQ 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 239 SGVYVGDPSKLSMKAAGKVMKLEETGGSIIIGGFKAIKREKSSPPKAPRDPRL- PKPKGQ 298
QY 286 TWGSPKGLRMLPRAISARLGSMTLKSITSEKGGYHLYETPEPEGVVSLQSRISV 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 299 TWGSPKGLRMLPRAISARLGSMTLKSITSEKGGYHLYETPEPEGVVSLQSRISV 358
QY 346 MTFVSYASNILRPLSYAADAALSNFYPRVAVNTVYPOAIRBERLV- DGLKFGQL 404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 359 LTPSYVVAADVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 418
QY 405 HPRTOGVETLGTIYSSSLFPMRAKGRVLLNLYIGGAKNPELISKTESOLVEVYVDRDLK 464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 419 HPRTOGVETLGTIYSSSLFPMRAKGRVLLNLYIGGAKNPELISKTESOLVEVYVDRDLK 478
QY 465 MLTKPKADPLVYVGRVWPAIQPFLVGHLDPLSTAKAMNDNGLEGLFLGNYVSVVAL 524
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 479 WIKPKDAPRPRVAVGRVWPAIQPFLVGHLDPLSTAKAMNDNGLEGLFLGNYVSVVAL 538
QY 525 GRVCGAVAVASEVTGFLSRVAYK 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 539 GRVCGAVAVASEVTGFLSRVAYK 562

```

RESULT 9
Q9TOP9 PRELIMINARY: PRT: 563 AA.

AC Q9TOP9
DI 01-MAY-2000 (TRMBLrel. 13, Created)
DI 01-MAY-2000 (TRMBLrel. 13, Last sequence update)
DI 01-MAY-2000 (TRMBLrel. 13, Last annotation update)
DE HERBICIDE-RESISTANT PROTOPORPHYRINOGEN OXIDASE PRECURSOR.
GN PPX1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaeae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-2674;
RX MEDLINE=99077317; PubMed=9862501;
RA Randalph-anderson B.L., Sato R., Johnson A.M., Harris E.H.,
RA Hauser C.R., Oeda K., Ishige F., Nishio S., Gillham N.W.,
RA Boynton J.E.;
RT "Isolation and characterization of a mutant protoporphyrinogen oxidase
RT gene from Chlamydomonas reinhardtii conferring resistance to porphyrin
RT herbicides."
RL Plant Mol. Biol. 38:839-858(1998).
DR EMBL: AF030179; AAC79630.1; -
FT TRANSIT PEPTIDE. 1
FT CHAIN 83
FT VARIANT 83 563 POTENTIAL.
FT OXIDASE. HERBICIDE-RESISTANT PROTOPORPHYRINOGEN
SQ SEQUENCE 563 AA; 59834 MW; 074584EC935CA3AF CRC64;

```

Query Match 53.7%; Score 1527.5; DB 10; Length 563;
Best Local Similarity 59.3%; Pred. No. 2.8e-104;
Matches 299; Conservative 74; Mismatches 114; Indels 17; Gaps 6;
QY 61 SSAVDGGPAE-----IDCVYVAGISGLCTAQVMSANY--PMLMTEARDRAG 107
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 60 AAAATGAPTAGAGAVAKLIDVWYVIVYVGGGLVYVGOALAAQHKIQNPLVTEARERVG 119
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 108 GNITTVERDGYLMEGRPSPDPMLTMAVDCGLKDDVLYLGPVNAPRFLYMKGKLRV 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 120 GNITVMSGDGYVMEGRPSPDPMLTMAVDCGLKDDVLYLGPVNAPRFLYMKGKLRV 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 168 SKLIDLPFFDLMSIPGKTRAGFGPGL- RSPSPGHEBSVDFQFVARNLNGEVEFRLLEPFC 226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 180 SGL-DALFFDLMSIPGKTRAGFGPGL- RSPSPGHEBSVDFQFVARNLNGEVEFRLLEPFC 238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 227 SGVYVGDPSKLSMKAAGKVMKLEETGGSIIIGGFKAIKREKSSPPKAPRDPRL- PKPKGQ 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

DB 239 GGVYAGDPSKLSMKAFAFNRIWILKEKNGSIVYGAIKLFPQROSNPAPRDPRLPKPKQ 298
 QY 286 TVGSPRKGRLRMLPAISARLGSKLSMKSITKSEKGYHLTYTPPGVYLSQSRIV 345
 DB 239 TVGSPRKGRLRMLPAISARLGSKLSMKSITKSEKGYHLTYTPPGVYLSQSRIV 345
 QY 346 MTPVSNNTLRPLSVAADALSNFYPPVAVTISYPOEAI RDRRLV-DEGLKGFQGL 404
 DB 359 LTAPEYVAADLVKEQADAAAALAGSFDYPMGAVLTLVPLSAVREERKAKSDGSDYVFGQL 418
 QY 405 HPRROGVTLTGTIYSSSLFPRAPKRVLLNTYTGAKRPELLISKRESDLYVNDRLR 464
 DB 419 HPRROGVTLTGTIYSSSLFPRAPKRVLLNTYTGAKRPELLISKRESDLYVNDRLR 478
 QY 465 MLIKRKAADPLVGVVWVWPAQIAPQFLVGHLDLTLSTAKAAMNDGELGFLGNGVYVAL 524
 DB 479 MWIKRDAKPRVGVVWVWVWPAQIAPQFLVGHLDLTLSTAKAAMNDGELGFLGNGVYVAL 538
 QY 525 GRCVGAATEVAEAVTGFELSKRAYK 548
 DB 539 KRVEHGYESAAANLAKSVSKAAVYK 562

RESULT 10
 Q9SLM5 PRELIMINARY; PRT; 502 AA.
 ID Q9SLM5 (TREMBlrel. 13, Created)
 AC Q9SLM5 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DR PROTOPORPHYRINOGEN IX OXIDASE.
 GN HEMG.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicot; Rosidae; eurosids I;
 OC Fabales; Fabaceae; Papilionoideae; Glycyne.
 OX NCBI_TaxID=3847;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20063348; PubMed=10594126;
 RA Kanjo N., Nishio S., Narita S., Oeda K., Inokuchi H.;
 RT "Nucleotide sequence of a cDNA clone encoding protoporphyrinogen IX
 oxidase (Accession No. AB025102) from soybean (Glycine max L.)."
 RL Plant Physiol. 121:1383-1383(1999).
 DR EMBL; AB025102; BAA76348.1; -
 DR InterPro: IPR000205; -
 DR InterPro: IPR000205; -
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR PRINTS: PR00419; ADXRDTASE.
 DR SEQUENCE 502 AA; 55046 MW; EFC8DF6EADCC3444 CRC64;

Query Match 18.4%; Score 522.5; DB 10; Length 502;
 Best Local Similarity 29.1%; Pred. No. 2.7e-30;
 Matches 148; Conservative 92; Mismatches 225; Indels 43; Gaps 13;

QY 61 SSAYDGGAAALDCVYAGISGICLAQVMSANYPNLMTVTEARDRAGNITTVEDDGYL 120
 DB 3 SSATDNDNRVSKRAVAVVAGVSGLAAYAKLSHGLDVTVEAEGRGRLRSVSDGDLW 62
 QY 121 EGGPNSFQSPDMLTMAVDC-GLKDDLVLDGDPNAPRFVLMKGLRVPVSKLIDLDFPFL 179
 DB 63 DEGANNTWESIEIKGLIDALGLOEKQPFISQHKRYIKNAPRLLVPRNPAALLKSKLL 122
 QY 180 SIPGKLARAGRPGLRPSPPGH-----EESVQVYRNLGEBERLIEPFGSGVYVD 233
 DB 123 SAOSKIHILIFEPFMKRSDDSNVCDENSYSVGRFPERFRHGVVYDYLDPVGGTSSAD 182
 QY 234 PSKLSMKAFAFGVYKLEBTEGSIIGGF--KAIKERSSTPKAPRDPRLPKPKGOTVGSFR 291
 DB 183 PESTLSMRHSFPELMLNLRKRGSTIAGALQSKLPAKREKTEGNNRTRALRKKRKGSGF-SFQ 241

QY 292 KGLRMLPDAISARLGSKLSMKSITKSEKGYHLTYTPPGVYLSQSRIV 345
 DB 242 GGMQTLTDTLCKELKDDLKLKNEVLTILAYGHOSSSQWMSITSNQSTQDWD--AVI 299
 QY 346 MTPVSNNTLRPLSVAADALSNFYPPVAVTISYPOEAI RDRRLV-DEGLKGFQGL 404
 DB 300 MTPVSNNTLRPLSVAADALSNFYPPVAVTISYPOEAI RDRRLV-DEGLKGFQGL 404
 QY 399 KCGFQHP-----RTQGVETLGTIYSSSLFPRAPKRVLLNTYTGAKRPELLISKRESDLY 455
 DB 349 EFGVYLVPSKREKNGKLTGLTFSSMMRPPRPSDLYLTYTFIGYQNNELMAGASDELR 408
 QY 456 EYVVDRLRMLIKRKAADPLVGVVWVWPAQIAPQFLVGHLDLTLSTAKAAMNDGELGFLG 515
 DB 409 KIVTSDLRKLL--GAEGETFVNHFWYMSKGFPLRGRNYGSVLAIDKIERD--LPGFFPA 464
 QY 516 GNVYSGVALRCVEGAYEVAEAVTGFELSKRAYK 543
 DB 465 GNYKGLSVYKRAIASGCAADLVITSLN 492

RESULT 11
 Q67814 PRELIMINARY; PRT; 436 AA.
 ID Q67814 (TREMBlrel. 07, Created)
 AC Q67814 (TREMBlrel. 07, Last sequence update)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE.
 GN HEMG.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OC NCBI_TaxID=63363;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham R.A., Overbeek R., Snead M.A., Keller M., Anjay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RT Nature 392:355-358(1998).
 DR EMBL; AE000768; AAC07778.1; -
 DR InterPro: IPR000205; -
 DR InterPro: IPR002937; -
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR SEQUENCE 436 AA; 48987 MW; EDSFZBIBRCDIDEBF7 CRC64;

Query Match 18.1%; Score 515; DB 2; Length 436;
 Best Local Similarity 29.9%; Pred. No. 7.7e-30;
 Matches 144; Conservative 103; Mismatches 170; Indels 64; Gaps 16;

QY 73 DCVYAGAGISGICLAQVMSANYPNLMTVTEARDRAGNITTVEDDGYLMEBGRPSQSDP 132
 DB 5 EYVYVAGISGISTAYRKLKREGVDVYVEKDDRIGGITHYVKEKGLYFVGAQTILADDE 64
 QY 133 MLTMAVDCGLKDDLVLDGDPNAP-RFVLMKGLRVPVSKLIDLDFPFLMSIPGKLARAGRP 191
 DB 65 VIDFLKEAGIEP--VEASPSKRYRVIYKKGRLIPLPMSVPEFLKTLTLLSIKTKLV-LTE 121
 QY 192 IGLRSPPGHEE--SVQGFVYRNLGEBERLIEPFGSGVYVDGPKLSMKAFAFGVYK 249
 DB 122 IFRK-----GVDDDISADPVEHNEGEBFLNVAAPFISGVYAGDPKLSIKATPKLYA 177
 QY 250 EFTGSIIGTETKA-IKERSSTPKAPRDPRLPKPKGOTVGSFRKRLMDDAISARL--- 305
 DB 178 OKKYGSLI---KAFIKKETA-----GPKKLI-SFEGEGELINMADQKLEVN 221
 QY 306 --GSKLTKSMKLSITKSEKGYHLTYTPPGVYLSQSRIVTAVSVAANTLRPLSVA 363

Db 222 TENVLRMR-KEEDEFRLDVG-----KKEVETKSVVAVASPAVTSYLLKEVFS 269
 QY 364 AADALSNFYPPVAVITISYPOEAI RDERLVDGELKGFQGLHPRGTGVEFTLGTYSLSLF 423
 Db 270 ASEBEFKIDYPPVAVVWG-----VEGKFPVRLSSSSGKRRILGAMFMSKLF 318
 QY 424 PNRAPKGRVLLANTIGAKNPEILSKTESQLEVVDRDLKMLIKPRAODPLV--VGVV 481
 Db 319 PGRAPQOKRELLTVLGGATREVEITLSEEEIENIVERLEKEIL-----QIDCIDFMHVOK 373

RESULT 12
 Q9FYV8 PRELIMINARY; PRT; 544 AA.
 AC Q9FYV8
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE PROTOPORPHRYRINOGEN IX OXIDASE.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
 OC Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Volrath S.L.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF273767; AAC00946.1; -
 SQ SEQUENCE 544 AA; 59020 MW; 101CEFS6EB9A151B CRC64;

Query Match 16.1%; Score 459; DB 10; Length 544;
 Best Local Similarity 28.3%; Pred. No. 1.5e-25;
 Matches 142; Conservative 88; Mismatches 226; Indels 46; Gaps 16;

QY 76 IVGAGISGLCIAOVMSANYPNLMTVEARDRAGNITTVERDGYLMEBGPNSFQPSDPMILT 135
 Db 50 VVAGAGVSGLAAYRILRQSGVAVVTFEADRAAGKIRTNSEGGHFWDEGAVTMEGEMEAS 109
 QY 136 MAV--DGLKXDDIVLGDNARRFVIMKGLRPPVSKLTDLPFRLDMSIPGKLRAGFGPI-- 192
 Db 110 RLDDDLGDDKQOYRPNQHKRYIVKDGAPRLIISDPISLMSKSSVLSSTKSKIALDFEFPFLY 169
 QY 193 ---GLRPPSPGHE---EVEQFVRRNLGGEVEFERLIEPFCSSGVYVGDPSKLSMKAAGF 245
 Db 170 KKANTRNSGKVSSEHLSSESGFCFRHFGREYVDYVDFVAGTSAQDPELSIRRAIFRA 229
 QY 246 VKLLEPFGSIIIGSTFKAIKERSSTPKARDPRLPKRGQVIG--SFRKGIKMLPDAISAR 304
 Db 230 LMMNERKRYGSAIVGAILSKLAKGDVKTTHDSSGKRMRNRFVSPFHGMQSLINALHNE 289
 QY 305 LG-SKLLKLSKLSITKS-----EKGGYHLLTYTPR--GVVSIQSR---SIVWTVVSYVA 353
 Db 290 VGDNDVAKLGEVLSLACTFEDVPALGRWSTISVDSKSGDKDLANSNOTFPAVINTAP---L 346
 QY 354 SNILRPL-----SVAADALSNFYPPVAVITISYPOEAI RDERLVDGELKGFQGLHP-- 406
 Db 347 SNVRBMMKFTGGAPVYLDLDFPKMDYLPILSMYTAAPKDDYK-----PLEBFGVLIIPYK 400
 QY 407 --RIGQVEFTLGTIVSSSLFPNRPAPKGRVLLANTIGAKNPEILSKTESQLEVVDRDLK 464
 Db 401 EQQKHGKTKTGTGTFSSMMFPDRAPODYLYTTFVGGSHNBDLAGAPSTILKQLVTSDLKK 460
 QY 465 MLIKPKAODPLVYGVVWVPAIQPLVGHLDLITSTAKAMN--DNGLEGLFLGNGVYSGVA 523

Db 461 LD--GVEGQPTFKHYVWGNAPF--LYGH-DVSSVLEALEIKNEKKMLPGFVYGNKDDGLA 515
 QY 524 LGRVCAEAVASEVTGFLSRY 545
 Db 516 VGSVIASGSKADLAISYLSH 537

RESULT 13
 Q9LYA7 PRELIMINARY; PRT; 501 AA.
 AC Q9LYA7
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE PROTOPORPHRYRINOGEN OXIDASE-LIKE PROTEIN.
 GN F18022_10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
 RA Rudd S., Lemcke K., Mayer K.F.X.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A1163817; CAB87761.1; -
 DR Interpro: IPR000205; -
 DR PRINMS: PR00419; ADXRDTASE.
 SQ SEQUENCE 501 AA; 54829 MW; 50CB8E1F3C0ED2F0 CRC64;

Query Match 16.0%; Score 456; DB 10; Length 501;
 Best Local Similarity 27.0%; Pred. No. 2.1e-25;
 Matches 142; Conservative 95; Mismatches 194; Indels 94; Gaps 18;

QY 63 AVDGGPAAELDCVIVGAGISGLCIAOVMSANYPNLMTVEARDRAGNITTVERDGYLME 122
 Db 13 AVSGKRVV-----VYAGAGVSGLAAYRILRQSGVAVVTFEADRAAGKIRTNSEGGH 67
 QY 123 GPNRQPSDPMI--TMAYVDCGLKDDIVLGDNARRFVIMKGLRPPVSKLTDLPFRLDMSI 181
 Db 68 GANTMTAEPEVGSLLDDDLGLREKQOFPISQKRYIVRNGVPMIPTNPIELVYSSVLSL 127
 QY 182 PKLRAGFGPI-----GLRPPSPGHESEVQFVRRNLGGEVEFERLIEPFCSSGVYVGDPSK 236
 Db 128 QSKPQILLEPFLMKKSSVSDASABESVSEFQNHFGQVYDYLIDPFGVGTSAADPS 187
 QY 237 LSMKAAFGKWKLEETGSSII-----GTFKAIKERSSTPKARDPRLPKRPGQ 285
 Db 188 LSMKHSFPLD--NSFGSIIYGAIKRTKFAAKGKSRDTPKSPGTKKSGRS----- 236
 QY 286 TVGSPRKGIRMLPDAI-----SARLSK--LKLS-----WKLSSITKSEKGGYHL 328
 Db 237 --FSRGGMQILPDLTKLSLSDHETINDSKVLSLVSNSGSRQEMWMSLSCVSHNE----- 288
 QY 329 TYTEPEGVVSLSRSIVMTVPSYVANSI--LRPLSVA---DALSNFYPPVAVIT 380
 Db 289 -----TQRN-----PHYDAAPLCNPKEMKVMKGGQFPQDINFLPILNYPILSVLI 333
 QY 381 ISYPOEAI RDERLVDGELKGFQGLHP--RIGQVEFTLGTIVSSSLFPNRPAPKGRVLLANTY 437
 Db 334 TFFTEKRYK-----RPLEGFGVLLIPRSKQKHGEFTLGTFLRSSMMFPDRSPSDVHLVYTF 387
 QY 438 IGGAKNPEILSKTESQLEVVDRDLKMLIKPRAODPLVYGVVWVPAIQPLVGHLDLIT 497
 Db 388 IGGSRNQLBLAKASIDELKQVYVTSIDLRL--GVBEPEVSNHYHYWRAKAFPLDYDSSY--DSV 444

Oy 498 STAKAANDNGLELFLGNGVYSGVALRCVCEGAYEVASVTFGL 542
 Db 445 MEALDKM-ENLDLPGFFYAGNHRRGLSVGKSIASGCKRADDLIVISYL 488

RESULT 14

OQ9FMS9 PRELIMINARY; PRT; 547 AA.

AC OQ9FMS9;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE PROTOPORPHYRINOGEN IX OXIDASE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=98162728; PubMed=9501997;
 RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
 RT Sequence features of the regions of 1,191,918 bp covered by seventeen
 RT physically assigned P1 clones."
 RL DNA Res. 4:401-414(1997).
 DR EMBL: AB007650; BAB08301.1; -;
 SQ SEQUENCE 547 AA; 60266 MW; 5474575A8B707A6E CRC64;

Query Match 16.0%; Score 456; DB 10; Length 547;
 Best Local Similarity 27.0%; Pred. No. 2,3e-23;
 Matches 142; Conservative 95; Mismatches 194; Indels 94; Gaps 18;

Oy 63 AVDGGPAEELDVCYIVGAGISGLCIAOVMSANYPNLMTVEARDRAGGNTTVERDGYLME 122
 Db 59 AVSGKRYA-----VVGAGVSGLAAYKIKSLGNLVYTFEADGRGKIRSYMONGLIMDE 113
 Oy 123 GPNFSPDPMI-TMAVDCGLKDDLVLDPNAPRFVLMKGLRVPVPSKLTLDLPEFLDMSI 181
 Db 114 GANFMTTEAPRPGVSLDLDLGRKQKQFPISQKRRIVANGVPMPLPNIPIELVTVSSVLSR 173
 Oy 182 PGLRAGGPI-----GLRPPRPHESVEQVFRNMLGGEVFERLIERPFGSGVYVGPBRK 236
 Db 174 QSKFQILLEPFLMKKSSKVSADASAESVSEFQRHFGQEVVDYLDLDFVGGTSAADPDS 233
 Oy 237 LSMKAAPGKVMKLEETGSSII-----GGTFPAIKERSSTPKARDPRLPKPKGQ 285
 Db 234 LSMKHSFQDLN--NSFSITVGAIRFKFAKGGKSRPTKSPGTTKSGRS----- 282
 Oy 286 TVGSRKGLRMLPDAI-----SARLGSK-LKLS-----WKLSSITKSEKGGYHL 328
 Db 283 --FSFKGGMQLLPLDLCKLSHDEINLDSKVLSTSYNSGSHQEWMSLSVSHNF----- 334
 Oy 329 TTEPREGVYSIQSRSTVMTVPSYVASN--LRPLSYAAA-----DALSNTYTPVAVT 380
 Db 335 -----TORON-----PHYDAARPLCNVKEKVMKGGQFQDNLNPLRETIYMPVSVLI 379
 Oy 381 ISYPOAIRDRERLVDGLKFGGOLHP---RTQGVETLGTIYSSSLFPNRAKGVLLNLY 437
 Db 380 TTFYKRYK-----RPLEGFGVILRPEQKHGKRTLGTLSMMMPFRPSRSDVHNLITTF 433
 Oy 438 IGGAKNPRLKSTESQVLEVVYDRDLRKMILRKAQDRPLVGVVWPAOAIPOFLVGHIDTL 497
 Db 434 IGGSRNOGLAAKSTDELKQVYVTDLQRL--GVGEPEVSNHYRKAFFPLDSSY-DSV 490
 Oy 498 STAKAANDNGLELFLGNGVYSGVALRCVCEGAYEVASVTFGL 542
 Db 491 MEALDKM-ENLDLPGFFYAGNHRRGLSVGKSIASGCKRADDLIVISYL 534

RESULT 15
 O64385 PRELIMINARY; PRT; 404 AA.

AC O64385;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4).
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BINTJIE;
 RA Johnston D.J., Droz E., Rochaix J.D., Malnoe P.C.;
 RL Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O.
 DR EMBL: AJ225108; GAA12401.1; -;
 DR Mendel; 29309; SOLtu;3077;29309.
 DR InterPro; IPR00205; -;
 KW Oxidoreductase.
 SQ SEQUENCE 404 AA; 44454 MW; A16C21D30AC46A34 CRC64;

Query Match 14.9%; Score 424; DB 10; Length 404;
 Best Local Similarity 28.7%; Pred. No. 3,5e-23;
 Matches 123; Conservative 83; Mismatches 152; Indels 70; Gaps 15;

Oy 60 PSSAVDGGPAEELDVCYIVGAGISGLCIAOVMSANYPNLMTVEARDRAGGNTTVER 115
 Db 3 PSAQED-----KONCKRKYAVIGAGVSGLAAYKIKHGNLVYTFEADGRGKIRSLISQ 57
 Oy 116 DGYLWEGNPSFQSDPMLTMAVD-CGLKDDLVLDPNAPRFVLMKGLRVPVPSKLTLDL 174
 Db 58 DGLIMDEGANVTMSESGVDYFLDLSLGRKQKQFPLSQNRKYIARNGPTLTPSPIDLI 117
 Oy 175 FPDLMSTPGLRAGGPIGLRPS-----PGHESVQVFRNMLGGEVFERLIERPFGSGVY 230
 Db 118 KSNFSLSTGSKLQMLFEPRLMKKNTLTKVSDHEHSVSGFQRHFGKEVVDYLDLDFVAGTC 177
 Oy 231 VQDPSKLSMKAAFGKVMKLEETGSSIIIGSTFKA---IKERSSTPKARDPRLPKPKGQ 286
 Db 178 GQDPSLSMHLSPRELMLNLEKRGSVYVGAIRSKLSPIKQKQP-----PKTVNKKRQ 232
 Oy 287 VGSFR--KGLRMLPDAI-----SARLGSK-LKLS-----SKLSSITKSEKGG 325
 Db 233 RGSFSLGGMQTLPLDAICDKLEDELRLNSRVLELSCSCSDAIDSMSIFSASPKRQA 292
 Oy 326 YHLTTEPREGVYSIQSRSTVMTVPSYVASNILRPLSYAANAADLNSFY---YPPVAVTI 381
 Db 293 EESFSD-----AVIMTAPLDCVKS--KIAKKNPNPLNFIPEVVDYVPLSVYTI 339
 Oy 382 SYPOAIRDRERLVDGLKFGGOLHP---RTQGVETLGTIYSSSLFPNRAKGVLLNLY 438
 Db 340 TFKKESVKH-----RLEGFGVILRPEQKHGKRTLGTLSMMMPFRDAPVWVLYLTFEV 393
 Oy 439 GGAKNPEI 446
 Db 394 GGSRRNEL 401

Search completed: July 3, 2001, 10:42:33
 Job time: 571 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:27:57 ; Search time 61.92 Seconds
(without alignments)
674.155 Million cell updates/sec

Title: US-09-508-418-2
Perfect score: 2844
Sequence: 1 MTTTPIANHPNIFTHQSSSS.....EGAYEVASEVTCFLSRYAYK 548

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2821	99.2	548	T04058	protoporphyrinogen
2	2565.5	90.2	557	T07116	protoporphyrinogen
3	2075	73.0	537	G85021	protoporphyrinogen
4	2061	72.5	545	T02005	protoporphyrinogen
5	315	18.1	436	B70473	protoporphyrinogen
6	505	17.8	504	T04076	protoporphyrinogen
7	485	17.1	470	D47045	coproporphyrinogen
8	456	16.0	501	T48595	protoporphyrinogen
9	424	14.9	404	T07121	protoporphyrinogen
10	407.5	14.3	467	D83800	protoporphyrinogen
11	368.5	13.0	477	S65684	protoporphyrinogen
12	363.5	12.8	462	A75435	protoporphyrinogen
13	362.5	12.7	477	A56449	protoporphyrinogen
14	362.5	12.7	477	S68367	protoporphyrinogen
15	317	11.1	468	T34679	probable protoporp
16	302.5	10.6	490	T36088	protoporphyrinogen
17	254	8.9	435	F86601	protoporphyrinogen
18	234	8.9	435	A81518	protoporphyrinogen
19	249	8.8	435	D72023	protoporphyrinogen
20	231.5	8.1	424	D81739	protoporphyrinogen
21	221	7.8	426	F70869	probable hemyr - M
22	211	7.4	424	A71476	probable protoporp
23	189	6.6	539	S50472	hypothetical prote
24	174.5	6.1	421	C75327	oxidoreductase, FA
25	169.5	6.0	411	A75361	hypothetical prote
26	153.5	5.4	416	T37023	probable oxidoredu
27	148	5.2	440	B71153	hypothetical prote
28	146	5.1	430	E84212	hypothetical prote
29	144	5.1	488	T47787	hypothetical prote

30	143	5.0	490	2	A84861	probable amine oxi
31	141	5.0	448	2	H70947	hypothetical prote
32	139	4.9	570	2	S54134	phytoene dehydrog
33	136.5	4.8	467	2	T35307	probable oxidoredu
34	134	4.7	499	2	T51346	monomine oxidase
35	132.5	4.7	527	2	S03974	amine oxidase (fla
36	131	4.6	566	2	F71403	hypothetical prote
37	127	4.5	574	2	A96612	hypothetical prote
38	125.5	4.4	472	2	S74886	phytoene dehydrog
39	123.5	4.3	526	2	JT0528	amine oxidase (fla
40	123	4.3	582	2	S29314	phytoene dehydrog
41	122.3	4.3	489	2	A47259	corticosteroid bin
42	122	4.3	471	2	S76290	hypothetical prote
43	121	4.3	404	2	A69467	hypothetical prote
44	119.5	4.2	472	2	T35936	probable dehydrog
45	119.5	4.2	478	2	A47693	putrescine oxidase

ALIGNMENTS

RESULT 1

T04058
Protoporphyrinogen oxidase (EC 1.3.3.4) IX precursor, chloroplast - common tobacco

C:Species: Nicotiana tabacum (common tobacco)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T04058

R:Lermontova, I.; Kruse, E.; Mock, H.P.; Grimm, B.
Proc. Natl. Acad. Sci. U.S.A. 94, 8895-8900, 1997

A:Title: Cloning and characterization of a plastidial and a mitochondrial isoform of t
A:Reference number: Z15186; MUID:97385200

A:Accession: T04058
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-548 <LEER>

A:Cross-references: EMBL:Y13465; NTD:92370332; PIDN:CAA73865.1; PID:92370333

A:Experimental source: Strain SK1
A:Gene: ppx1

C:Function:
A:Pathway: tetrapyrrole synthesis

C:Keywords: chloroplast; oxidoreductase
E:1-20/Domain: transit peptide (chloroplast) #status predicted <TNP>

F:51-548/Product: protoporphyrinogen oxidase IX #status predicted <MAT>

Query Match	Best Local Similarity	Score	2821;	DB 2;	Length	548;
Matches 544; Conservative 1; Mismatches 3; Indels 0; Gaps 0;						
1	MTTTPIANHPNIFTHQSSSSPLAFINRTSFTPEFSSISKRNSVNCNGWTRKCSVAKDYTP	60				
1	MTTTPIANHPNIFTHQSSSSPLAFINRTSFTPEFSSISKRNSVNCNGWTRKCSVAKDYTP	60				
61	SSAVYGGPAAEIDCVYVAGISGLCIAQVMSANYNLWVTRARROAGNITTVBERDGLW	120				
61	SSAVYGGPAAEIDCVYVAGISGLCIAQVMSANYNLWVTRARROAGNITTVBERDGLW	120				
121	EGSPNSFOPSDPMLTMAVDCGLKDDLVLDGPNAPRFVLMKGRKIRVPSKLTLDLFFDILMS	180				
121	EGSPNSFOPSDPMLTMAVDCGLKDDLVLDGPNAPRFVLMKGRKIRVPSKLTLDLFFDILMS	180				
181	IEGKLRAGEPPIGLRPPSPGHEEVEQFVRRNLTGGEVERLIEPFCSGVYVAGDPSKLSMK	240				
181	IEGKLRAGEPPIGLRPPSPGHEEVEQFVRRNLTGGEVERLIEPFCSGVYVAGDPSKLSMK	240				
241	AAPGKVMKLEFEGSSITIGTTPFAIKERSSTPKAPRPPPLPKGGTGVGSRKGLRMLPDA	300				
241	AAPGKVMKLEFEGSSITIGTTPFAIKERSSTPKAPRPPPLPKGGTGVGSRKGLRMLPDA	300				
301	ISARLGSKTLKLSMTKLSITKSEKGYHLTYETPEGVVLSQSSIVMTPVSVVANSILRPL	360				
301	ISARLGSKTLKLSMTKLSITKSEKGYHLTYETPEGVVLSQSSIVMTPVSVVANSILRPL	360				

OY 361 SVAADALSNFYPPVAVVTISYPOEAI RDRERLVDGELKGFQGLHPRPQGVETLGTIYSS 420
 |||||||
 Db 361 SVAADALSNFYPPVAVVTISYPOEAI RDRERLVDGELKGFQGLHPRPQGVETLGTIYSS 420
 OY 421 SLFPNRARPKGAVLLNNTYIGAKNPEIISKTESQLEVVDRDLRKMLIKPKAODPLVYGV 480
 |||||||
 Db 421 SLFPNRARPKGAVLLNNTYIGAKNPEIISKTESQLEVVDRDLRKMLIKPKAODPLVYGV 480
 OY 481 VMPQAIPOFLVGHLDLTLSTAKAAMNDGELFLFGNNVSGVALGRCVEGAYEVAASEVTG 540
 |||||||
 Db 481 VMPQAIPOFLVGHLDLTLSTAKAAMNDGELFLFGNNVSGVALGRCVEGAYEVAASEVTG 540
 OY 541 FLSTRYAYK 548
 |||||||
 Db 541 FLSTRYAYK 548

RESULT 2
 T07116
 protoporphyrinogen oxidase (EC 1.3.3.4) IX, chloroplast - potato
 C:Species: Solanum tuberosum (potato)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
 C:Accession: J07116
 R:Johmston, D.J.
 submitted to the EMBL Data Library, April 1998
 A:Reference number: Z15932
 A:Accession: T07116
 A:status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-557 <J0H>
 A:Cross-references: EMBL:AJ225107; NID:e1289157; PIDN:CAA12400.1; PID:e1289158
 A:Experimental source: cv. Bintje
 C:Genetics:
 A:Genome: nuclear
 C:Function:
 A:Pathway: tetrapyrrole biosynthesis
 C:Keywords: chloroplast; oxidoreductase

Query Match 90.2% Score 2565.5; DB 2; Length 557;
 Best Local Similarity 88.9%; Pred. No. 1.5e-181;
 Matches 499; Conservative 21; Mismatches 32; Indels 9; Gaps 2;
 OY 1 MTTPIANHPNIFTHQ-----SSSSP-LAFLNRTSFIPISSISKRNSVNCNGMTRC 51
 |||||
 Db 1 MTTPIANHPNIFTHRSRPLRSPSSSSSPFLFLNRTNFIPIRSTSRNSVNCNGMTRC 60
 OY 52 SVAKDYIVPSSAVDGGPAELDCVYVAGISGLCIAQVMSANPNLWTEARDRAGCNT 111
 |||||
 Db 61 SVAKDYIVPPEVVDGQFPELDCVYVAGISGLCIAKVIANSYPNLWTEARDRAGCNT 120
 OY 112 TVERDGYLMEGGRPKSPQSPDMLTMAVDCGLKDDLVLDGDNAPRFVLMKGLRVPVSKLT 171
 |||||||
 Db 121 TVERDGYLMEGGRPKSPQSPDMLTMAVDCGLKDDLVLDGDPARFVLMKGLRVPVSKLT 180
 OY 172 DLPEFDLMSIPGKLRAGFGPIGLRPPSPGHESEVEQFVRRNLGGEVFERLIEPFCSSVY 231
 |||||||
 Db 181 DLPEFDLMSIPGKLRAGFGAIIRPSPGHESEVEQFVRRNLGGEVFERLIEPFCSSVYA 240
 OY 232 GDRPSKLSMKARFQVW 291
 |||||||
 Db 241 GDRPSKLSMKARFQVW 300
 OY 292 KGLRMLPDAISARLGSKTKLWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVW 351
 |||||||
 Db 301 KGLRMLPDAISARLGSKTKLWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVW 360
 OY 352 VASNIIRPLSVAADALSNFYPPVAVVTISYPOEAI RDRERLVDGELKGFQGLHPRPQGV 411
 |||||||
 Db 361 VASNIIRPLSVAADALSNFYPPVAVVTISYPOEAI RDRERLVDGELKGFQGLHPRPQGV 420
 OY 412 ETLGTTIYSSSLFPNRARPKGAVLLNNTYIGAKNPEIISKTESQLEVVDRDLRKMLIKPKA 471

Db 421 ETLGTTIYSSSLFPNRARPKGAVLLNNTYIGAKNPEIISKTESQLEVVDRDLRKMLIKPKA 480
 OY 472 QDPLVGVVRRWPAIQFLVGHLDLTLSTAKAAMNDGELFLFGNNVSGVALGRCVEGA 531
 |||||||
 Db 481 QDPLVGVVRRWPAIQFLVGHLDLTLSTAKAAMNDGELFLFGNNVSGVALGRCVEGA 540
 OY 532 YEVAASEVTGFLSTRYAYK 548
 |||||||
 Db 541 YEVAASEVTGFLSTRYAYK 557

RESULT 3
 G85021
 protoporphyrinogen oxidase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G85021
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: AB5001; MIDID:20083488
 A:Accession: G85021
 A:status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-537 <STO>
 A:Cross-references: GB:INC_001268; NID:g7268212; PIDN:CAB77739.1; GSPDB:GN00140
 C:genetics:
 A:gene: AT4g01690
 A:map position: 4

Query Match 73.0% Score 2075; DB 2; Length 537;
 Best Local Similarity 75.4%; Pred. No. 2.4e-145;
 Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;
 OY 22 LAFLNRTSFIPISSISKRNSVNCNGW--RTRCSYAKDYVPSAVDGGPAEL--DCVI 76
 |||||
 Db 3 LSLRPTTQSLPLPSPKPN-IRLNYKKPLRLKCSVAAGGPTVSSKIBEGGGTTITDLCVI 61
 OY 77 VGAGISGLCIAQVMSANPNLWTEARDRAGCNTTYVERDGYLMEGGRPKSPQSDP 132
 |||||||
 Db 62 VGGGISGLCIAQVMSANPNLWTEARDRAGCNTTYVERDGYLMEGGRPKSPQSDP 121
 OY 133 MTLMAVDCGLKDDLVLDGDNAPRFVLMKGLRVPVSKLTLDPFLDMSIPGKLRAGFGPI 192
 |||||||
 Db 122 MTLMAVDCGLKDDLVLDGDPARFVLMKGLRVPVSKLTLDPFLDMSIPGKLRAGFGAL 181
 OY 193 GDRPSPGHESEVEQFVRRNLGGEVFERLIEPFCSSVYVWVWVWVWVWVWVWVWVWVWVW 252
 |||||||
 Db 182 GDRPSPGHESEVEQFVRRNLGGEVFERLIEPFCSSVYVWVWVWVWVWVWVWVWVWVWVW 241
 OY 253 GGSITGTFKAIKESRNPAPRDRPKPKQVYVWVWVWVWVWVWVWVWVWVWVWVWVWVW 312
 |||||||
 Db 242 GGSITGTFKAIKESRNPAPRDRPKPKQVYVWVWVWVWVWVWVWVWVWVWVWVWVW 301
 OY 313 WKLSTTKSEKGYHLTYETTPGCVVSLQSRSTVMTVPVSVASNIIRPLSVAADALSNFY 372
 |||||||
 Db 302 WKLSTTKSEKGYHLTYETTPGCVVSLQSRSTVMTVPVSVASNIIRPLSVAADALSNFY 361
 OY 373 YPPVAVVTISYPOEAI RDRERLVDGELKGFQGLHPRPQGVETLGTIYSSSLFPNRARPK 432
 |||||||
 Db 362 YPPVAVVTISYPOEAI RDRERLVDGELKGFQGLHPRPQGVETLGTIYSSSLFPNRARPK 421
 OY 433 LLNNTYIGAKNPEIISKTESQLEVVDRDLRKMLIKPKAODPLVYGVVWVWVWVWVWVWVW 492
 |||||||
 Db 422 LLNNTYIGAKNPEIISKTESQLEVVDRDLRKMLIKPKAODPLVYGVVWVWVWVWVWVWVW 481
 OY 493 HLDLTLSTAKAAMNDGELFLFGNNVSGVALGRCVEGAYEVAASEVTGFLSTRYAYK 548
 |||||||
 Db 482 HLDLTLSTAKAAMNDGELFLFGNNVSGVALGRCVEGAYEVAASEVTGFLSTRYAYK 537

RESULT 4

T02005

protoporphyrinogen oxidase (EC 1.3.3.4) - Arabidopsis thaliana

N:Alternate names: protein T15B16.13

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 21-Jul-2000

C/Accession: T02005; J05488

R:Stonking, T.; SmLth, R.

A:Submitted to the EMBL Data Library, November 1998

A:Description: The sequence of A. thaliana T15B16.

A:Reference number: Z14488

A:Accession: T02005

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-545 <STO>

A:Cross-references: EMBL:AF104919; NID:g3859590; PID:g3859604

A:Experimental source: cultivar Columbia

R:Harita, S.; Tanaka, R.; Ito, T.; Okada, K.; Taketani, S.; Inokuchi, H.

A>Title: Molecular cloning and characterization of a cDNA that encodes protoporphyrinoge

A:Reference number: J05488; MUID:97136707

A:Accession: J05488

A:Molecule type: mRNA

A:Residues: 1-217,226-545 <NAR>

A:Cross-references: DBJ:D83139; NID:g1183006; PIDN:BA11820.1; PID:g1877018

C:Genetics:

A:Map position: 4

A:Intons: 126/3; 188/3; 225/1; 274/2; 301/2; 358/3; 412/1; 446/3

A:Note: T15B16.13

C:Function:

A:Description: catalyzes the removal of six electrons from protoporphyrinogen to generat

A:Pathway: chlorophyll biosynthesis; heme biosynthesis

C:Keywords: oxidoreductase

Query Match 72.5%; Score 2061; DB 2; Length 545;

Best Local Similarity 74.3%; Pred. No. 2,6e-144; Indels 18; Gaps 5;

Matches 404; Conservative 49; Mismatches 73;

Db 1 22 IAFNRTSFPSPSSSRKRNVSNCNGM--RTRCSVAKDYVPSSAVDGGPAEL--DCVY 76

Db 3 LSILRPTQSGLPFSFKPN-IRLNVYKPLRLRCSVAGGPTVSSKIEGGGTTITTDVYI 61

Db 77 VAGISGLCIAQVMSANY---PNLWVTEARDRAGNITTTVERDGYLWEEGPNFQSDP 132

Db 62 VGGISGLCIAQVMSANY---PNLWVTEARDRAGNITTTVERDGYLWEEGPNFQSDP 121

Db 133 MLMAVNDGKIDLVLDGPPNAPRFLVLMKGRVPRKLDLPDFDLMSIPGKLRAGGFI 192

Db 122 MLTIVVDSGLKDDVLADDPAPRFLVLMKGRVPRKLDLPDFDLMSIPGKLRAGGFI 181

Db 193 GLRPPSPGHEEVEQFVRNMLGVEVRELIERPFCG-----VYVGDPSKLSMKAARQ 244

Db 182 GIRPSPGHEEVEQFVRNMLGVEVRELIERPFCG-----VYVGDPSKLSMKAARQ 241

Db 245 KVMKLEETGSSITGGTFPKAKERSSTPKAPRDRPKPKQVYVGSFRKGLMIPDAISAR 304

Db 242 KVMLKLENGSISIIIGTFPKAOERKNAPKAKERDRPKPKQVYVGSFRKGLMIPDAISAR 301

Db 305 LGSKLKSWKLSITKSEKGGYHLYETPREGVYVLSQRSITVMTPVPSVANSILRPLSVA 364

Db 302 LGSKVKLSWLSITKSEKGGYHLYETPREGVYVLSQRSITVMTPVPSVANSILRPLSVA 361

Db 365 ADALSNFYPPVGVAVTISYPOEARLDRERLVYDGLKGFQGLHPRGVEVTGLTIYSSELFP 424

Db 362 ANLSKLYYPRVAAYSTSYPKKAIRTECCLDGLKGFQGLHPRGVEVTGLTIYSSELFP 421

Db 425 NRPAPKGRVLLANTIGAKNPEIILSKTESQLVYVYVDRDLKMLIKPKAODPLVYGVAVWPQ 484

Db 422 NRPAPKGRVLLANTIGAKNPEIILSKTESQLVYVYVDRDLKMLIKPKAODPLVYGVAVWPQ 481

Db 485 AIPQFLVGHLDLTSTFAKAAVNDNGLGFLGNYVYSVAGLRCVPEGAVYASEVYTFSLR 544

Db 482 AIPQFLVGHLDLITFAKSSITSGYBELFLGNYVYVAGLRCVPEGAVYAEIVANNPMSR 541

Db 545 YAYK 548

Db 542 YAYK 545

RESULT 5

B70473

protoporphyrinogen oxidase - Aquifex aeolicus

C:Species: Aquifex aeolicus

C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C/Accession: B70473

R:Deckerl, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

A:Reference number: A70300; MUID:98196666

A:Accession: B70473

A:Molecule type: DNA

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Cross-references: GB:AE000768; NID:g2984249; PIDN:AA007778.1; PID:g2984251; GB:AE00

A:Experimental source: strain VF5

C:Genetics:

A:Gene: hemo

Query Match 18.1%; Score 515; DB 2; Length 436;

Best Local Similarity 29.9%; Pred. No. 2,5e-30;

Matches 144; Conservative 103; Mismatches 170; Indels 64; Gaps 16;

Db 73 DCVIVGAGISGLCIAQVMSANYPNLWVTEARDRAGNITTTVERDGYLWEEGPNFQSDP 132

Db 5 EYVYVAGISGLSTAYRLKKEGVDPVYVYKDRIGGTHYKKEKGYLFEVGAQITLADQE 64

Db 133 MLTMAVNDGKIDLVLDGPPNAPRFLVLMKGRVPRKLDLPDFDLMSIPGKLRAGGFI 191

Db 65 YIDFLPKAEKIEP--VEASPSKRYRIRYIKKGRILRPLPMSVPELTKTPLLSLTKTKLV--LTE 121

Db 192 IGLRPPSPGHEE--SVEQFVRNMLGVEVRELIERPFCG--VYVGDPSKLSMKAARQYKWL 249

Db 122 IFKR-----GVDDSDIADPVRHGEPLRNVYVAFISGVVAGDPEKLSLKHATPKLYEA 177

Db 250 EFTGSSITGGTFPKA-IKERSSTPKAPRDRPKPKQVYVGSFRKGLMIPDAISAR-- 305

Db 178 OKKYGSLI----KAFIKRKA-----GPGKGLI-SFGEGELGILVALAOKLEVH 221

Db 306 --GSKLTKSWKLSITKSEKGGYHLYETPREGVYVLSQRSITVMTPVPSVANSILRPLSVA 363

Db 222 TENVVLRRK-KEDEFRLDVRG-----KVEYTKSVVAVASPATSSYLKKEYSFS 269

Db 364 AADALSNFYPPVGVAVTISYPOEARLDRERLVYDGLKGFQGLHPRGVEVTGLTIYSSELFP 423

Db 270 ASEEFDKLIDYRPPVVYVNVG-----VEGRPKYRLLSSGKKRIILGAMFMKSLF 318

Db 424 PRRAPKGRVLLANTIGAKNPEIILSKTESQLVYVYVDRDLKMLIKPKAODPLV--GVYV 481

Db 319 PGRADQKELLIVFLVFGGATDREIVLSEELIENIYERELKELI-----QIICIDIFMNHQK 373

Db 482 WQAIPTQFLVGHLDLTSTFAKAAVNDNGLGFLGNYVYSVAGLRCVPEGAVYASEVYTF 541

Db 374 WRRAIPTYLLIDYRFLINLAQENKED--YRGLFLGNWMLYGVSTADCLRASKRVAKOKLSLF 431

Db 542 L 542

Db 432 L 432

RESULT 6

T04076

protoporphyrinogen oxidase (EC 1.3.3.4) IX, mitochondrial [validated] - common tobacc

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
 C:Accession: R04076
 R:Riemontova, I.; Kruse, E.; Mock, H.P.; Grimm, B.
 Proc. Natl. Acad. Sci. U.S.A. 94, 8895-8900, 1997
 A:Title: Cloning and characterization of a plastidial and a mitochondrial isoform of tobacco
 A:Reference number: 215186; MUID:97385200
 A:Accession: R04076
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-504 <LEB>
 A:Cross-references: EMBL:Y13466; NID:92370334; PIDN:CAA73866.1; PID:92370335
 A:Experimental source: strain SRI
 C:Genetics:
 A:Gene: ppXII
 C:Function:
 A:Description: EC 1.3.3.4 [validated; MUID:97385200]
 A:Pathway: tetrapyrrole synthesis
 C:Keywords: mitochondrion; oxidoreductase

Query Match 17.8%; Score 505; DB 2; Length 504;

Best Local Similarity 26.8%; Pred. No. 1.7e-29; Indels 60; Gaps 17;

Matches 149; Conservative 100; Mismatches 208;

60 PSSAVDGPAAELDCVIVAGISGICIAQVMSANVPMVMTVEARDRAGNTTVERDGYL 119

3 PSAGDHRKSSMK-RVAIVGAGVSGIAAAYKIKHGLNVTVPEAGKAGKGRISQDGLI 61

120 WEQGNFQSPDPMVMAVD-CGLKDLVLDGPNAPRFVLMKGLRVPYKLDLPPFDL 178

62 WDEGNTMTESEGDVTFELIDSLGIREKQFPPLSONKRYIARANGTPVLLPSPDILIKSNF 121

179 MSITGKLRAGSGPI-----GLRSPRPGHESEVDFVRRNIGVEFELIERFGCGVYVGD 233

122 LSTGSKLMLLEPIIKMKKILSIQVSDH-ESVSGFQRHREKEVVDYLDIDPFVAGTGGGD 180

224 PSLKSMKAQFKVWKLLETTGSIIGTFKA-IKERSSTPKAPRDPRIKPKGQTVSGFRK 292

181 PDSLSMHHSPFELMNLKRFKRSYVILGAIKRSKLNKPKGQPRPKTSANKKRQRFSEFLG 240

293 GLRMLPRAI-----SARLGSK-LKL-----SKLSSITSEKGGYHLYTEFP 333

241 GMQTLTDAICDLREDELRLMSRVLELSCCTEDSADISWISIASHPKRSSESESD-- 298

334 EGVVSLGSRISVMTVPSVVAANIIRPLSVAAADALSNFY-----YRPVGAVTISPOFAIR 389

299 -----AVIMTAPLCDVMSM--KIAKKGNPPLMFIPEVDYVPLSVYITTFKRENVK 347

390 DERLVDELKGFQJHP--RTQGVETLGTIYSSSLPKNRPAKGRVLLMVIIGKAKNPEI 446

348 -----YPLEGFGVLYPFSKEQHGKLTGLTFSSMMPDRAPNPNVLYLTFVGGSRREL 401

447 LSKTESQVLEVVDLDRKMLIKRPAQDPRLVYGVVWQALPQFLVGH-LDTLSTAKAAMN 505

402 AKARTELEKTELVTDLKQLD--GAEGEPYVYVNHLYWSKAPP--LYGHNVYSVDLDAIDKME 457

506 DNGLEGLFLGNGVYVVALGRCVGEVAVSEVTFGL 542

458 KN-LPGLFYAGNHRGGLSVGKALSSGCMADLVISTYL 493

RESULT 7

D47045

coproporphyrinogen III oxidase / coproporphyrinogen IX oxidase hemY - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000

C:Accession: D47045; D69640

R:Hansson, M.; Hederstedt, L.

J. Bacteriol. 174, 8081-8093, 1992

A:Title: Cloning and characterization of the Bacillus subtilis hemBY gene cluster, which

A:Reference number: A47045; MUID:93094140

A:Contents: 3518

A:Accession: D47045

A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-470 <HAN>
 A:Cross-references: GB:M97208; NID:9143041; PIDN:AAA22519.1; PID:9143045
 A:Note: Sequence extracted from NCBI backbone (NCBIN:119989, NCBIPI:119993)
 R:Kunze, F.; Ogasawara, N.; Mosszer, I.; Albertoni, A.M.; Alloni, G.; Azavedo, V.; Ber
 C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capano, V.; Careri, N.M.;
 A.: Ehrlich, S.D.; Emmerson, P.T.; Entlan, K.D.; Errington, J.; Fabret, C.; Ferrati,
 Nature 390, 249-256, 1997
 A:Authors: Fouldier, C.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galluzzi, A.; Gal
 lech, J.; Harwood, C.R.; Henaut, A.; Hillbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
 Koester, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 Y, M.; Ogasawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scani
 A:Authors: Schleich, S.; Schroeter, R.; Scifione, F.; Sekiguchi, J.; Sekowska, A.; Se
 akuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchida
 T.; Winters, P.; Wipart, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zumbstein, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:98044033
 A:Accession: D69640
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-470 <KUN>
 A:Cross-references: GB:Z99109; GB:AL009126; NID:92633260; PIDN:CAB12854.1; PID:e11830
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: hemY

Query Match 17.1%; Score 485; DB 2; Length 470;

Best Local Similarity 28.7%; Pred. No. 4.6e-28; Indels 42; Gaps 15;

Matches 139; Conservative 98; Mismatches 206;

75 VIVAGISGICIA-----QVMSANVP-NIMVTVEARDRAGNTTVERDGYLMEQPNFSF 127

9 VIIIGGITGMAAFYWEKELKKNLPLELIVVAVSRVGGKIQYVKKQGIYIERRPSDFL 68

128 --QSPDPMVMAVDGKLDVLDGPNAPRFVLM-----KKIRPYPSKLDLPPFDL 177

69 ERKKSAPQVVK-DLGL-ELLVNNATGQSYVLVNRTLHPMKGAVMGIPTKIAFVSTG 125

178 IMSIPGKLRAGSGPIGLRSPRPGHESEVDFVRRNIGVEFELIERFGCGVYVGD 237

126 LPSLSKARAAMDFT-LPASKTRKDDSGEFRRKRYGDEVENLEPDLISGTYAGDITKL 184

238 SMKAQFKVWKLLETTGSIIGTFKAIKERSSTPKAPRDPRIKPKGQTVSGFRKGLRML 297

185 SLKSTPQGYQTEQKHKRSLIDM-----KTRRQSGGQQLFAKKQGG-FQTLSTGLQTL 237

298 PDAISARLQSKLKLSSITKSEKGYHLYTEFPREGVYSIQSISYMTVPSVYASNIL 357

238 VEIEKQD--KLTKYKYGKVTKLKSHSGCVSLELDNG-VTLDAQSYVTVLPHKAAAGML 294

358 RPLSVAAADALSNFYVYVAVTISYPOFAIRDERLRYVDELKGFQJLHPRPQGVETLGTI 417

295 SEL--PAISHLKMMHSTAVANALGFPEGSVQMEH-----EGTGFVYSRNSDPAITACT 346

418 YSSSLPKNRPAKGRVLLMVIIGKAKNPEILSKTESQVLEVVDLDRKMLIKRPAQDPRLV 477

347 WTKKWRPHAAPEKTLRLRAYVGGKADDESIVLSDNDIINIVLEDKKVM--NINDEPMT 404

478 CVRWVQDALPQFLVGHNDLSTAKAAMNDGELGFLGNGVYVVALGRCVGEVAVSEVTFGL 537

405 CVTRHMSMPQYVNHGHKORIKELREAL-ASAVYGVYMTGASFGVGIPICDIQKRAAVSD 463

538 VTGFL 542

464 ALLYL 468

RESULT 8

Db 326 RDNHL-----HGQVAADBEIGIKATVTSKAKHGRAPREGHVLLRVFFKDDLPQKALA 378
 QY 449 KTESQLEVEVDDLRKMLIKPKRAQ-DPLVYGVVWVPOAIRPFLIPDLTSLTAKAAMND 507
 Db 379 EAVHEVARLRFQ-----AQQDPLMHAAYADMRGNKRNPAQVGHLDHIAIRKALPRN 427
 QY 508 GLEGEFLGNGVYVGVALGRCVEGAYEVAEVEVTFGL 542
 Db 428 ----IQVAGSSYTGVDIPDCVNAAGRTAARDVVAAL 458

RESULT 13

A56449
 protoporphyrinogen oxidase (EC 1.3.3.4) - human
 C:Species: Homo sapiens (man)
 C:Date: 27-Oct-1995 #sequence_revision 27-Oct-1995 #text_change 20-Jun-2000
 C:Accession: J04971; A56449; I38995
 R:Publ. H.; Kobrea, A.M.; Rosipal, R.; Nordmann, Y.; Deybach, J.C.
 Biochem. Biophys. Res. Commun. 226, 226-230, 1996
 A:Title: Protoporphyrinogen oxidase: Complete genomic sequence and polymorphisms in the
 A:Reference number: J04971; MUID:96400242
 A:Accession: J04971
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-477 <P>
 A:Cross-references: EMBL:X99450
 R:Nishimura, K.; Takekani, S.; Inokuchi, H.
 J. Biol. Chem. 270, 8076-8080, 1995
 A:Title: Cloning of a human cDNA for protoporphyrinogen oxidase by complementation in vi
 A:Reference number: A56449; MUID:95229621
 A:Accession: A56449
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-477 <N>
 A:Cross-references: GB:D38537; NID:g854663; PIDN:BA07538.1; PID:g854664
 R:Dalley, T.A.; Meissner, P.; Dalley, H.A.
 J. Biol. Chem. 269, 813-815, 1994
 A:Title: Expression of a cloned protoporphyrinogen oxidase.
 A:Reference number: A49858; MUID:94117488
 A:Accession: I38995
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-477 <R>
 A:Cross-references: EMBL:U26446; NID:9837327; PIDN:AAA67690.1; PID:9837328
 C:Comment: This enzyme acts as the penultimate step in the heme biosynthetic pathway, at
 C:Genetics:
 A:Gene: GDB:PPOX; PPO
 A:Cross-references: GDB:118852; OMTM:600923
 A:Map position: 1q23-1q23
 A:Introns: 29/3; 74/3; 113/2; 157/3; 206/1; 269/3; 290/1; 329/3; 366/3; 416/3; 431/1
 C:Superfamily: phytoene dehydrogenase
 C:Keywords: oxidoreductase

Query Match 12.8%; Score 363.5; DB 2; Length 477;

Best Local Similarity 25.5%; Pred. No. 4.4e-19; Matches 138; Conservative 83; Mismatches 170; Indels 151; Gaps 23;

QY 75 VIVGAGISGLCIAQVMSANY-----PMLMTEARDRAGNITTFE-RDGYLMEGPN 125
 Db 6 VIVGGISGLI-----AASVHLSRAPCPKPVYLVESSEKLGWIRSVGKPNDAIPELPR 59
 QY 126 SHQPSDPM-----LTMAYDGGKJDDL--VLGDPNAP--RFLVMKGLRVPVSKLTDLPFPD 177
 Db 60 GTRPAGALDARFLLLVSELDSEVLVPRGDHPAQNRRFLYVGGALHMLPGLRGL---- 115
 QY 178 LMSIPGKLKRAAGFPIGLRPPSP-----GHE--ESVDFVRRNNGEVEF 218
 Db 116 -----LRSPSPSKPLFMWAGLRELLTKRKGKPEPDTVHSFAQRRLGPEVA 159
 QY 219 EKLIEPFGCVYVDPSPSKLMSKAAFGKVMKLEETGGSIIGTFPAIKERSSTPKARDPDR 278

Db 160 SLAMDSLCRGVAFAGNSRSLRSCFSLFQAEDVHRSILGLLGA-----GR 207
 QY 279 LRPKGGQTV-----GSFRKGLMRLDPAISARLGS-----LKL----- 312
 Db 208 TPQDSDALIRKALAEKRSQMSLRGLMLPQALFETHLTSIGVSVLRQPCVGLSLQAEGR 267
 QY 313 WKLSSTKRSKGGYHLTEPEEGVSLQSRSIYMTVPSYVANSILRPLSVAAADALSNEY 372
 Db 268 MKV-SLRDS-----SLEADHVISAIPASVLSLELPEAARPLARLALSAT 310
 QY 373 YRPVGAVTISYPPALIDDERLVDGELKFGOLHPRQGVETLGTIYSSSLFPNR--APK 430
 Db 311 AVSAVAVNLQYOGANL-----PVQFGHLVPSSEDPVGLGVYDVSVAFPQDQSPG 362
 QY 431 -RVLVLYVIGGAKRPETLSKTESQLEVEVDRDLRKMILIKPKADPPLV--VGVHVP---- 483
 Db 363 LRVTVM--LGS-----SWLQTLERASGCVLSQELFQORADQAAATQGLKEMPSHCL 411
 QY 484 ----QAIPOFLVGHLDPLTSTAKAAMNDNGLEGLFEGNYSVGVALGRCVEGAYEVAEVEV 538
 Db 412 VHLKNCIPQYTLGHMQKLESARQFLTAHRLPFLAGASY--EGVAVNDCTESGQAASV 470
 QY 539 TG 540
 Db 471 LG 472

RESULT 14

S68367
 protoporphyrinogen oxidase (EC 1.3.3.4) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 19-May-2000
 C:Accession: S68367; S65630
 R:Dalley, T.A.; Dalley, H.A.; Meissner, P.; Prasad, A.R.K.
 Arch. Biochem. Biophys. 324, 379-384, 1995
 A:Title: Cloning, sequence, and expression of mouse protoporphyrinogen oxidase.
 A:Reference number: S68367; MUID:96132666
 A:Accession: S68367
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-477 <DR>
 A:Cross-references: EMBL:U25114; NID:g793939; PIDN:AAA96003.1; PID:g793940
 R:Takekani, S.; Yoshinaga, T.; Furukawa, T.; Kohno, H.; Tokunaga, R.; Nishimura, K.;
 Eur. J. Biochem. 230, 760-765, 1995
 A:Title: Induction of terminal enzymes for heme biosynthesis during differentiation o
 A:Reference number: S65629; MUID:95331315
 A:Accession: S65630
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 2-14;164-178 <TAK>
 C:Genetics:
 A:Genome: nuclear
 C:Superfamily: phytoene dehydrogenase
 C:Keywords: mitochondrion; oxidoreductase

Query Match 12.7%; Score 362.5; DB 2; Length 477;

Best Local Similarity 24.8%; Pred. No. 5.2e-19; Matches 133; Conservative 76; Mismatches 188; Indels 139; Gaps 19;

QY 75 VIVGAGISGLCIAQVMSANY-----PMLMTEARDRAGNITTFE-RDGYLMEGPN 131
 Db 6 VIVGGISGLAASVHLSRAPCPKPVYLVESSEKRLGWRISRSDDAIFELGFRGIRPAP 65
 QY 132 PM-----LTMAYDGGKJDDL--VLGDPNAP--RFLVMKGLRVPVSKLTDLPFPDLSI 183
 Db 66 ALGARTLLVSELDSEVLVPRGDHPAQNRRFLYVGGTLPHPSPSGLRGL----- 115
 QY 184 KLRAGFPIGLRPPSP-----GHE--ESVDFVRRNNGEVEF 224
 Db 116 -----LRSPSPSKPLFMWAGLRELLTKRKGKPEPDTVHSFAQRRLGPEVAISLAMD 165

OY 225 FCGSYVVDPEPKLSMKAFAFGVWMLKLEETGSIIGTFEKAIKERSSTPKARPRDLPKPKG 284
 Db 166 LCRVGFANNSNELSTRSGFPLFQAEOGTRSTILGLLGGQS-----PPDS 213
 OY 285 QTV-----GSFRKGLRMLPDATSARLGSK-----KLS---WKLSSI 318
 Db 214 SLIRQARAERMSQMSLRNGGLEVLFPQALHNHLSKGVYVLSGQPCVGLDPEGMKYSL- 272
 OY 319 TKSEKGYHLTYETPEGVYLSQSRSYWTVPSYVAANTLRPLSVAAADALSNFPPYVGA 378
 Db 273 -----GSSSLEAHHTISALPASELSKLLPAAEAPLRITLSTIKAVSAV 316
 OY 379 VTISYPOAIRDERLVDELKFGQLHPRTOGVETLGTIYSSSLFPNR---APKRVLL 435
 Db 317 VNLQYRGNACL-----PVGPFHLVPSSEDPVYLGIVYDVSVAFPEDGMPSELRYTVM 368
 OY 436 NYIGGAKNPBELLSTESQVLEVDRLRKMILKPKAD-----PLVGVVWPQ 484
 Db 369 --LGG-----YWLQKLRKAAGHOLSPELFOOAOEAAATQGLKPEPCHLVHLKN 417
 OY 485 AIPQFLVGHLDLTLSTAKAAMNDNGLEGLFGGNVYSGVALGRVGEVAYEVAEYVVG 540
 Db 418 CIPQTTIGHMOKLDSAMQFLTAQRPLTLTAGASY-EGVAVNDCTESGRQAAVAVLG 472

RESULT 15

T34679
 probable protoporphyrinogen oxidase - Streptomyces coelicolor (fragment)
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 03-Nov-2000
 C:Accession: T34679
 R:Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z21553
 A:Accession: T34679
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-468 <HAR>
 A:Cross-references: EMBL:AL023517; PIDN:CAAL18975.1; GSPDB:GNO0070; SCOE0B:SC1B5_01
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOE0B:SC1B5_01
 C:Superfamily: phyloene dehydrogenase

Query Match 11.1% Score 317; DB 2; Length 468;
 Best Local Similarity 26.5% Pred. No. 1.2e-15;
 Matches 129; Conservative 77; Mismatches 217; Indels 64; Gaps 17;

OY 81 ISGLCIAQVMSANYPNLMVTEARDRAGGNTTVERDGYLWEEGPNSEFQSPDP-MLTMAVD 139
 Db 1 IAGLAHAHRLEAGARVTVLEASGVGKLLPGEIAGVAVDLAGAEMLARREPAVGLARA 60
 OY 140 CGLKDDLVLDPPNAPRYLW-KGKLRVYPSKLLDLPFDLMSIPKLRAGRPITG----- 193
 Db 61 AGLADR--IQPPSTATASIMTRGALRPMPKG-----HWGAVPGTAAALSGVLSIEGLA 111
 OY 194 -----LRPPPGHESYEQFVVRNNGEVEFRLTEPFCGQVYVVDPSKLSMKAARQKV 246
 Db 112 RIERDAELPREVDGDDVAAGVYVAARLGRVYDRVLEPLGLGVAAGDAYRISLRSVAVPQL 171
 OY 247 WKLEBTGSSIIIGTFKAIKERSSTPKARPRDLPKPKGQTVGSPKGLRMLP-----DAIS 302
 Db 172 FEARHTHSLEAV-RMLQGRTA-----SPPSGPEVMKIEGGIGTLPPAVADSVR 221
 OY 303 ARLGSKLKLWKLSSITKSEKGYHLTYETPEGVYLSQSRSYWTVPSYVAANTLRPLSV 362
 Db 222 AR--GGEIYTOPAPVTELRRRTASDGWRIV---AGDRVLHAGAVVYVAVPAGPAAEILLREAP 276
 OY 363 AAAADALSNFYPPVGNATYISPOAIRDERLVDELKFGQLHPRTOGVETLGTIYSSSL 422
 Db 277 AAAAELSAVEASVALVTLAIR---RSEAAALPEGSGF--LVPPVVG---HTI-KAST 325

OY 423 FPNR-----APKGVLLNNTYGAKNPBELLSTESQVLEVVVDRLRKMILKPKAODPL 475
 Db 326 FASRKKGMWLADEDDPDLVYLRTSVGRYGDTEILLGRDAGLVAVSRHDLAEMTGLTAA--PV 383
 OY 476 VGVVWVPOAIRPQFLVGHLDLTLSTAKAAMNDNGLEGLFGGNVYSGVALGRVGEVAYEVA 535
 Db 384 AFRVTRMQDGLPQYPVGH--HARVARVRHHVAKLPLGLAVCAAYDGVGIPASIASAVAAA 441
 OY 536 SEVTFGL 542
 Db 442 DQIRGDL 448

Search completed: July 3, 2001, 10:34:11
 Job time: 374 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:00:01 ; Search time 49.77 Seconds
(without alignments)
667.509 Million cell updates/sec

Title: US-09-508-418-2

Perfect score: 2844

Sequence: 1 MTTPTTANPNRHTFHSSSS.....EGAYEVASVETGFLSRAYK 548

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- 1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:**
- 2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:**
- 3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:**
- 4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:**
- 5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:**
- 6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:**
- 7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:**
- 8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:**
- 9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:**
- 10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:**
- 11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:**
- 12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:**
- 13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:**
- 14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:**
- 15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:**
- 16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:**
- 17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:**
- 18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:**
- 19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:**
- 20: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:**
- 21: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:**
- 22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2836	99.7	548	AAV05200	N. tabacum protopo
2	2836	99.7	548	AAV05201	N. tabacum protopo
3	2113	74.3	539	AAW41608	Soybean protox-I.
4	2113	74.3	539	AAW41609	Cotton protox-I.
5	2113	74.3	539	AAW25740	Cotton protoporphy
6	2113	74.3	539	AAW25740	Cotton protoporphy
7	2108	74.1	539	AAW72921	Cotton protoporphy
8	2104	74.0	539	AAW72922	Cotton protoporphy
9	2078	73.1	537	AAW72923	Arabidopsis protop
10	2075	73.0	537	AAW90295	Protein having pro
11	2075	73.0	537	AAW41603	Arabidopsis thalial

Result	Score	Query Match	Length	ID	Description
12	2075	73.0	537	AAW25746	Arabidopsis protop
13	2075	73.0	537	AAW51347	Arabidopsis protop
14	2075	73.0	537	AAW42385	Arabidopsis thalial
15	2075	73.0	537	AAW72901	Arabidopsis protop
16	2074	72.9	537	AAW72924	Arabidopsis protop
17	2066	72.6	537	AAW72925	Arabidopsis protop
18	2055	72.3	543	AAW25739	Soybean protoporphy
19	2055	72.3	543	AAW92498	Glycine max protop
20	2055	72.3	543	AAV54100	The chloroplast-ty
21	2055	72.3	543	AAW72906	Soybean protoporphy
22	2006.5	70.6	536	AAW41611	Rape protox-1. Br
23	2006.5	70.6	536	AAW25742	Rape protoporphyri
24	2006.5	70.6	536	AAW25742	Oilseed rape protop
25	1987	69.9	528	AAW41607	Wheat protox-1. T
26	1987	69.9	528	AAW25738	Wheat protoporphyr
27	1987	69.9	528	AAW72903	Wheat protoporphyr
28	1978	69.5	481	AAW41605	Maize protox-1. Z
29	1978	69.5	481	AAW25745	Maize protoporphyr
30	1978	69.5	481	AAW51255	Maize protop-porphy
31	1978	69.5	481	AAW72903	Maize protoporphyr
32	1972	69.3	483	AAW90297	Maize protox-1 par
33	1972	69.3	481	AAW72926	Maize protoporphyr
34	1968	69.2	481	AAW72920	Maize protoporphyr
35	1933.5	68.0	560	AAW41610	Sugar beet protox-
36	1933.5	68.0	560	AAW25741	Sugarbeet protopor
37	1933.5	68.0	560	AAW72908	Sugar beet protop
38	1823	64.1	440	AAW41613	Sorghum protox-1.
39	1823	64.1	440	AAW25744	Sorghum protoporphy
40	1823	64.1	440	AAW72911	Sorghum protoporphy
41	1786	62.8	416	AAW42386	Arabidopsis thalial
42	1772	62.3	413	AAW42387	Arabidopsis thalial
43	1530.5	53.8	563	AAW49538	Chlamydomonas rein
44	1530.5	53.8	563	AAV54101	An algal protoporphy
45	1261	44.3	312	AAW41612	Rice protox-1. Or

ALIGNMENTS

Result	Score	Query Match	Length	ID	Description
1	2836	99.7	548	AAV05200	N. tabacum protopo
AC	2836	99.7	548	AAV05200	N. tabacum protopo
XX	2836	99.7	548	AAV05200	N. tabacum protopo
DT	2113	74.3	539	AAW41608	Soybean protox-I.
DE	2113	74.3	539	AAW41609	Cotton protox-I.
DE	2113	74.3	539	AAW25740	Cotton protoporphy
XX	2113	74.3	539	AAW25740	Cotton protoporphy
KW	2108	74.1	539	AAW72921	Cotton protoporphy
OS	2108	74.1	539	AAW72921	Cotton protoporphy
PN	2104	74.0	539	AAW72922	Cotton protoporphy
XX	2104	74.0	539	AAW72922	Cotton protoporphy
PD	2078	73.1	537	AAW72923	Arabidopsis protop
XX	2078	73.1	537	AAW72923	Arabidopsis protop
PF	2075	73.0	537	AAW90295	Protein having pro
XX	2075	73.0	537	AAW90295	Protein having pro
PA	2075	73.0	537	AAW41603	Arabidopsis thalial
XX	2075	73.0	537	AAW41603	Arabidopsis thalial
11-SEP-1997	97JP-0265084				
10-SEP-1998	98WO-JP04064				
18-MAR-1999					
WO9913087-A1					
Nicotiana tabacum.					
Hirooka T, Horikoshi M, Mametsuka K;					
WPI: 1999-215067/18.					
N-PSDB: AAX28278.					
(NIPPON NOYAKU CO LTD.					
Protoporphyrinogen oxidase; light-requiring herbicide tolerance; enzyme.					
Protoporphyrinogen oxidase tolerant to light-requiring herbicides, useful for constructing plants with such properties					
Claim 1: Page 34-39; 56pp; Japanese.					

CC This sequence is the Nicotiana tabacum protoporphyrinogen oxidase
 CC of the invention, which is tolerant to light-requiring herbicides. The
 CC enzyme and its derivatives can be applied to construct plants with high
 CC tolerance to light-requiring herbicides, useful in agriculture.
 XX

SQ Sequence 548 AA;

Query Match 99.7%; Score 2836; DB 20; Length 548;
 Best Local Similarity 99.8%; Pred. No. 8.7e-264;
 Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTTPTPIANHPNIFTHQSSSPFLAFLNRTSFIPESSISKRSVNCNGMRTKCSVAKDYTP 60
 | mttcplianhpnifthgssspflaflnrtsfipessiskrsvncngmrtkcsvakydtytp
 Db 1 mttcplianhpnifthgssspflaflnrtsfipessiskrsvncngmrtkcsvakydtytp 60
 QY 61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNIMLWTEARDRAGNITTVERRDGYLW 120
 | ssaavdggpaaeldcvivyagisglciaqvmsanypnimlwteardragnitvverrdgylw
 Db 61 ssaavdggpaaeldcvivyagisglciaqvmsanypnimlwteardragnitvverrdgylw 120
 QY 121 EEGPNSFQPSDPMITMAVDCGLKDDLVLDGPNAPRFVLMKGLRVPVSKLTDLPFDLMS 180
 | eegpnsfqpsdpmitmaavdcgkddlvldgpnaprflvmkglrvpvsykltdlpfdlms
 Db 121 eegpnsfqpsdpmitmaavdcgkddlvldgpnaprflvmkglrvpvsykltdlpfdlms 180
 QY 181 IPGKLRAGFGFGLRPPSPGHEESVEQFVVRNMGGEVEFERLIEPFCSGYVYVGDPSKLSMK 240
 | ipgklragfgfglrppspgheesveqfvvrnmggeveferliepfcsgyvyvgdpsklsmk
 Db 181 ipgklragfgfglrppspgheesveqfvvrnmggeveferliepfcsgyvyvgdpsklsmk 240
 QY 241 AAFGKVMKLEETGSGIIGTFKAIKERSSTPKAPRDLRPPKGGQTVGSRKGLRMLPDA 300
 | aafgkvmkleetgsgiiigtfkaikersstpkaprdlrppkggqtvgsrkglrmlpda
 Db 241 aafgkvmkleetgsgiiigtfkaikersstpkaprdlrppkggqtvgsrkglrmlpda 300
 QY 301 ISARLGSKTKLSMKLSSITKSEKGYHLTYETPEBQVVSLOSRSIWMTPVSYVASNILRPL 360
 | isarlgsktklsmklssttksekgyhlyetpebqvvslosrsiwmtpvsvyvasnilrpl
 Db 301 isarlgsktklsmklssttksekgyhlyetpebqvvslosrsiwmtpvsvyvasnilrpl 360
 QY 361 SVAADALSNFYPPYVAVTISYPOEARIDDERLVDSGLKGFQDLHPRPTQGVFTLGTIYSS 420
 | svaadalSNFYPPYVAVTISYPOEARIDDERLVDSGLKGFQDLHPRPTQGVFTLGTIYSS
 Db 361 svaadaalSNFYPPYVAVTISYPOEARIDDERLVDSGLKGFQDLHPRPTQGVFTLGTIYSS 420
 QY 421 SLFPNRPKRGVLLLNLYIGGAKNPELISKTESQLVVEVDRDLRKKMLIKPKAQPDLVVGVR 480
 | slfprnrpkrgvlllnlyiggaknpeLISKTESQLVVEVDRDLRKKMLIKPKAQPDLVVGVR
 Db 421 slfprnrpkrgvlllnlyiggaknpeLISKTESQLVVEVDRDLRKKMLIKPKAQPDLVVGVR 480
 QY 481 VMPQALPQFLVGHDLTSTAKAAMNDNGLEGLFLGGNYVSGVALGRCEVAGAYEVAASEVTG 540
 | vmpqalpqlvghdlstakaaMNDNGLEGLFLGGNYVSGVALGRCEVAGAYEVAASEVTG
 Db 481 vmpqalpqlvghdlstakaaMNDNGLEGLFLGGNYVSGVALGRCEVAGAYEVAASEVTG 540
 QY 541 FLSRYAYK 548
 | fLSRYAYK 548
 Db 541 fLSRYAYK 548

RESULT 2
 AA05201
 ID AA05201 standard; Protein: 548 AA.
 XX
 AC AA05201;
 XX
 DE 16-JUN-1999 (first entry)
 XX
 KW N. tabacum protoporphyrinogen oxidase mutant.
 XX
 OS Nicotiana tabacum.
 XX
 PN W09913087-A1.
 XX
 PD 18-MAR-1999.
 XX
 PF 10-SEP-1998; 98W0-JP04064.
 XX

PR 11-SEP-1997; 97JP-0265084.
 XX
 XX (NIPY) NIPPON NOYAKU CO LTD.
 PA
 XX
 XX Hirooka T, Horikoshi M, Mametsuka K;
 PI
 XX
 XX WPI: 1999-215067/18.
 DR
 DR N-PSDB: AAX28279.
 PR
 PR protoporphyrinogen oxidase tolerant to light-requiring herbicides,
 PI useful for constructing plants with such properties
 PS Example 5; Page 44-49; 56pp; Japanese.
 CC This sequence is the Nicotiana tabacum protoporphyrinogen oxidase
 CC of the invention, which is tolerant to light-requiring herbicides. The
 CC enzyme and its derivatives can be applied to construct plants with high
 CC tolerance to light-requiring herbicides, useful in agriculture.
 XX

Query Match 99.7%; Score 2836; DB 20; Length 548;
 Best Local Similarity 99.8%; Pred. No. 8.7e-264;
 Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTTPTPIANHPNIFTHQSSSPFLAFLNRTSFIPESSISKRSVNCNGMRTKCSVAKDYTP 60
 | mttcplianhpnifthgssspflaflnrtsfipessiskrsvncngmrtkcsvakydtytp
 Db 1 mttcplianhpnifthgssspflaflnrtsfipessiskrsvncngmrtkcsvakydtytp 60
 QY 61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNIMLWTEARDRAGNITTVERRDGYLW 120
 | ssaavdggpaaeldcvivyagisglciaqvmsanypnimlwteardragnitvverrdgylw
 Db 61 ssaavdggpaaeldcvivyagisglciaqvmsanypnimlwteardragnitvverrdgylw 120
 QY 121 EEGPNSFQPSDPMITMAVDCGLKDDLVLDGPNAPRFVLMKGLRVPVSKLTDLPFDLMS 180
 | eegpnsfqpsdpmitmaavdcgkddlvldgpnaprflvmkglrvpvsykltdlpfdlms
 Db 121 eegpnsfqpsdpmitmaavdcgkddlvldgpnaprflvmkglrvpvsykltdlpfdlms 180
 QY 181 IPGKLRAGFGFGLRPPSPGHEESVEQFVVRNMGGEVEFERLIEPFCSGYVYVGDPSKLSMK 240
 | ipgklragfgfglrppspgheesveqfvvrnmggeveferliepfcsgyvyvgdpsklsmk
 Db 181 ipgklragfgfglrppspgheesveqfvvrnmggeveferliepfcsgyvyvgdpsklsmk 240
 QY 241 AAFGKVMKLEETGSGIIGTFKAIKERSSTPKAPRDLRPPKGGQTVGSRKGLRMLPDA 300
 | aafgkvmkleetgsgiiigtfkaikersstpkaprdlrppkggqtvgsrkglrmlpda
 Db 241 aafgkvmkleetgsgiiigtfkaikersstpkaprdlrppkggqtvgsrkglrmlpda 300
 QY 301 ISARLGSKTKLSMKLSSITKSEKGYHLTYETPEBQVVSLOSRSIWMTPVSYVASNILRPL 360
 | isarlgsktklsmklssttksekgyhlyetpebqvvslosrsiwmtpvsvyvasnilrpl
 Db 301 isarlgsktklsmklssttksekgyhlyetpebqvvslosrsiwmtpvsvyvasnilrpl 360
 QY 361 SVAADALSNFYPPYVAVTISYPOEARIDDERLVDSGLKGFQDLHPRPTQGVFTLGTIYSS 420
 | svaadalSNFYPPYVAVTISYPOEARIDDERLVDSGLKGFQDLHPRPTQGVFTLGTIYSS
 Db 361 svaadaalSNFYPPYVAVTISYPOEARIDDERLVDSGLKGFQDLHPRPTQGVFTLGTIYSS 420
 QY 421 SLFPNRPKRGVLLLNLYIGGAKNPELISKTESQLVVEVDRDLRKKMLIKPKAQPDLVVGVR 480
 | slfprnrpkrgvlllnlyiggaknpeLISKTESQLVVEVDRDLRKKMLIKPKAQPDLVVGVR
 Db 421 slfprnrpkrgvlllnlyiggaknpeLISKTESQLVVEVDRDLRKKMLIKPKAQPDLVVGVR 480
 QY 481 VMPQALPQFLVGHDLTSTAKAAMNDNGLEGLFLGGNYVSGVALGRCEVAGAYEVAASEVTG 540
 | vmpqalpqlvghdlstakaaMNDNGLEGLFLGGNYVSGVALGRCEVAGAYEVAASEVTG
 Db 481 vmpqalpqlvghdlstakaaMNDNGLEGLFLGGNYVSGVALGRCEVAGAYEVAASEVTG 540
 QY 541 FLSRYAYK 548
 | fLSRYAYK 548
 Db 541 fLSRYAYK 548

RESULT 3
 AA041608
 ID AA041608 standard; Protein: 539 AA.
 XX
 AC AA041608;
 XX

```

DN 20-APR-1998 (first entry)
XX Soybean protox-1.
DE
KW Protoporphyrinogen oxidase-1; protox-1; promoter; soybean;
KW herbicide resistance; breeding programme; probe; gene isolation;
KW genomic mapping.
XX Glycine max.
XX Key Location/Qualifiers
FH 55..1686
FT /*tag= a
FT /product= protox-1
XX MO9732028-A1.
XX 04-SEP-1997.
XX 27-FEB-1997; 97WO-US03343.
XX 21-JUN-1996; 96US-0020003.
XX 28-FEB-1996; 96US-0012705.
XX 28-FEB-1996; 96US-0013612.
XX (NOVS ) NOVARTIS AG.
XX Johnson MA, Volrath SL, Ward ER;
XX WPI: 1997-489209/45.
XX N-PSDB: AAV04309.
XX DNA containing a plant proto-porphyrinogen oxidase gene promoter -
XX optionally linked to a heterologous gene, especially to express
XX herbicide-resistant enzymes, and plants containing such constructs
XX Claim 33; Pages 68-71; 114pp; English.
XX
XX The present sequence is soybean protoporphyrinogen oxidase-1
XX (protox-1).
XX The protox-1 promoter can be used to express herbicide resistant
XX enzymes, specifically protox, i.e. a plant tissue, plant or progeny
XX containing a chimeric gene of the promoter and a heterologous
XX coding sequence. The plant can also be used in breeding programmes.
XX Also hybridising fragments of the protox coding sequence can be
XX used as probes, e.g. to isolate related genes or for genomic
XX mapping.
XX
XX Sequence 539 AA:
SQ

```

```

DB 274 kgqlvgsfkkglumlpealanslgnvklswklssitklynggnlftetpegmwslsgr 333
OY 343 SIWVWVSYVAANIIRPLSVAADALSNFYPPVGAVRTISYPOEAI RPERLVDGELKGRG 402
DB 334 svmltphvasnllhpllaadaalsqfypvasvvtvsykeatirkeclldgelkqfg 393
OY 403 QLHPRTOGVEETLGTIYSSSLFPNRAKGRVLLNNTIGAKNPELISKTESQLVEVDRDL 462
DB 394 qlhprsqgletlgtlyssslfponrapsgrvlllnylygatalcglstkegeylveaavrdl 453
OY 463 RKMILTRPKADPLVGVRRVPAQIPOPFLVGHDLTSTAKAAMNDGELRFLFGGNVYSGV 522
DB 454 rkmlhpnakdpvlvgyrvwpkaiptflvghdlldskamallrdsqfnglflgnvysgv 513
OY 523 ALGRVGEAYEVAEVTGFLSRYAYK 548
DB 514 algrvgeayevaevkelfsqyayk 539

```

RESULT 4

```

AAW41609
ID AAW41609 standard; Protein; 539 AA.
XX
AC AAW41609.
XX
DT 20-APR-1998 (first entry)
DE Cotton protox-1.
XX
XX Protoporphyrinogen oxidase-1; protox-1; promoter; cotton;
KW herbicide resistance; breeding programme; probe; gene isolation;
KW genomic mapping.
XX Gossypium hirsutum.
XX
XX MO9732028-A1.
XX
XX 04-SEP-1997.
XX
XX 27-FEB-1997; 97WO-US03343.
XX
XX 21-JUN-1996; 96US-0020003.
XX 28-FEB-1996; 96US-0012705.
XX 28-FEB-1996; 96US-0013612.
XX (NOVS ) NOVARTIS AG.
XX Johnson MA, Volrath SL, Ward ER;
XX WPI: 1997-489209/45.
XX N-PSDB: AAV04313.
XX
XX DNA containing a plant proto-porphyrinogen oxidase gene promoter -
XX optionally linked to a heterologous gene, especially to express
XX herbicide-resistant enzymes, and plants containing such constructs
XX Claim 35; Pages 80-82; 114pp; English.
XX
XX The present sequence is cotton protoporphyrinogen oxidase-1
XX (protox-1).
XX The protox-1 promoter can be used to express herbicide resistant
XX enzymes, specifically protox, i.e. a plant tissue, plant or progeny
XX containing a chimeric gene of the promoter and a heterologous
XX coding sequence. The plant can also be used in breeding programmes.
XX Also hybridising fragments of the protox coding sequence can be
XX used as probes, e.g. to isolate related genes or for genomic
XX mapping.
XX
XX Sequence 539 AA:
SQ

```

```

Query Match 74.3%; Score 2113; DB 18; Length 539;
Best Local Similarity 79.2%; Pred. No. 2.8e-194;
Matches 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;

```

Matches	401; Conservative	46; Mismatches	55; Indels	4; Gaps	1;
QY	47	WRTRGSYAKDYTVSSAVDGGPAAEELDCVTVGAGISGLCIAQVMSANY----	PNLMWTEA	102	
Db	34	fklrcslaeapltssakldgessladvvgglsqldaglakckhdvasnvttea	93		
QY	103	RDRAGSNITTVVERDGYLWEEGPNSEFQSDPMLTMAYDCCGLKDDLVLDPNAPRFVLMKKG	162		
Db	94	rdrtvgnilttverdyjlweegpnsfqpdpdltmavdsqldkdlvlqdpnaprfvlwvqk	153		
QY	163	LRPVPSKLTLDLPEFDLMSIRPKRAGFRIQLRPPSPGHESEVDFQVRNLDGGEVFERLI	222		
Db	154	lrpvpkskptldlpfidmslsgklrsgfsgalrpppyveesveeltrnldgaevferfi	213		
QY	223	EPPGSGVYVGDPEPSKLSMKAFAFGVWKLERTGSIIGTEFKAIKERSTPKAPRDRPLRP	282		
Db	214	epfcsgvyagdpdsksmkaafgrvwxkleeigsllgqtkltqernktpkpprdplpkp	273		
QY	283	KGGTVGSPFRKGLRMLDPAISARLGSKLKLSMKSITKSEKGGYHLTVETPBGVVSLSGR	342		
Db	274	kgqtvgsfkkgltmlpealanslgnvklswklsstlkgngynlftfclpbgmvsisgr	333		
QY	343	SIWMTVPSVYASNILRPLSYAADAALSNFYPPVGAVITSYQEAIRDRERLVDSGLKGF	402		
Db	334	svwmtlpsvhasnllhplsaadaalsqfyppvasvtvsypkkairkceclldgelkqfg	393		
QY	403	QLHPRTOGVEETLGTIYSSSLFPNRAKPGRVLLNTYIGAKNPEILSKTESQLVEVVRDL	462		
Db	394	qlhprsqgletlgtlyssslfnpnrapsgrvlllnyigagatngllsktegelveavdrdl	453		
QY	463	RKMLIKPKAODPLVGVVRWPOAIPQFLVGHLDLSTAKAANNDSLEGLFSGVYVSGV	522		
Db	454	rkmllnpnkdplvlgvrvwvkaipqflvghldlldsakmalrdsqfnglflgnyvsvv	513		
QY	523	ALGRCEGAYEVAASEVTGFLSRAVK	548		
Db	514	algrcvegayevaaevkeltlsgayk	539		
RESULT	5				
AAW25740		standard; Protein: 539 AA.			
AAW25740:					
01-MAR-1998		(first entry)			
Cotton protoporphyrinogen oxidase (protox-1).					
Protox-1; protoporphyrinogen oxidase; inhibitor; cotton;					
herbicide tolerance; herbicide resistance; transgenic plant.					
Gossypium hirsutum L.					
Key		Location/Qualifiers			
Misc-difference 365		/note- "substitution of Pro-365 by another amino acid, especially Ser, provides a modified protox tolerant to a herbicide (Claims 51-52)"			
Misc-difference 428		/note- "substitution of Tyr-428 by another amino acid, especially Cys or Arg, provides a modified protox tolerant to a herbicide (Claims 53-54)"			
W09732011-A1.					
04-SEP-1997.					
27-FEB-1997;		97WO-US03313.			
* 21-JUN-1996;		96US-0020003.			

PR	28-FEB-1996;	96US-0012705.			
PR	28-FEB-1996;	96US-0013612.			
PA	(NOVS)	NOVARTIS AG.			
PI	Heifetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER;				
DR	WP1: 1997-448683/41.				
DR	N-FSDB; AAT86123.				
PT	New DNA encoding plant protoporphyrinogen oxidase enzyme - and herbicide resistant mutants, useful to prepare plants resistant to herbicide which therefore kills undesired vegetation only				
XX	Claim 7; Page 139-141; 196pp; English.				
PS	This protein comprises cotton protoporphyrinogen oxidase (protox-1), an enzyme that catalyzes the oxidation of protoporphyrinogen IX to protoporphyrin IX. Its amino acid sequence was deduced from an cotedon Protox-1 cDNA clone (see AAT86123). Sites within the cotton Protox-1 gene have been identified that can be mutated to encode a modified protox that is resistant to protox inhibitors and hence tolerant of certain herbicides. Plants, especially crop plants, may be engineered for resistance to protox inhibitors via mutation of the native protox gene to a resistant form, or they may be transformed with a gene encoding an inhibitor-resistant form of a plant protox enzyme, such as claimed forms from wheat, soybean, cotton, sugarcane, oilseed rape, rice and sorghum (see AAW25738-48). Application of herbicide will then kill undesired vegetation only. Protox enzymes can also be expressed in transformed host cells and used to identify inhibitors of protox enzyme activity, i.e. herbicide candidates, or to design herbicide tolerant forms of the enzyme.				
CC	Sequence	539 AA;			
CC	Query Match	74.3%; Score 2113; DB 18; Length 539;			
CC	Best Local Similarity	79.2%; Pred. No. 2, 8e-194;			
CC	Matches	401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;			
QY	47	WRTRGSYAKDYTVSSAVDGGPAAEELDCVTVGAGISGLCIAQVMSANY----	PNLMWTEA	102	
Db	34	fklrcslaeapltssakldgessladvvgglsqldaglakckhdvasnvttea	93		
QY	103	RDRAGSNITTVVERDGYLWEEGPNSEFQSDPMLTMAYDCCGLKDDLVLDPNAPRFVLMKKG	162		
Db	94	rdrtvgnilttverdyjlweegpnsfqpdpdltmavdsqldkdlvlqdpnaprfvlwvqk	153		
QY	163	LRPVPSKLTLDLPEFDLMSIRPKRAGFRIQLRPPSPGHESEVDFQVRNLDGGEVFERLI	222		
Db	154	lrpvpkskptldlpfidmslsgklrsgfsgalrpppyveesveeltrnldgaevferfi	213		
QY	223	EPPGSGVYVGDPEPSKLSMKAFAFGVWKLERTGSIIGTEFKAIKERSTPKAPRDRPLRP	282		
Db	214	epfcsgvyagdpdsksmkaafgrvwxkleeigsllgqtkltqernktpkpprdplpkp	273		
QY	283	KGGTVGSPFRKGLRMLDPAISARLGSKLKLSMKSITKSEKGGYHLTVETPBGVVSLSGR	342		
Db	274	kgqtvgsfkkgltmlpealanslgnvklswklsstlkgngynlftfclpbgmvsisgr	333		
QY	343	SIWMTVPSVYASNILRPLSYAADAALSNFYPPVGAVITSYQEAIRDRERLVDSGLKGF	402		
Db	334	svwmtlpsvhasnllhplsaadaalsqfyppvasvtvsypkkairkceclldgelkqfg	393		
QY	403	QLHPRTOGVEETLGTIYSSSLFPNRAKPGRVLLNTYIGAKNPEILSKTESQLVEVVRDL	462		
Db	394	qlhprsqgletlgtlyssslfnpnrapsgrvlllnyigagatngllsktegelveavdrdl	453		
QY	463	RKMLIKPKAODPLVGVVRWPOAIPQFLVGHLDLSTAKAANNDSLEGLFSGVYVSGV	522		
Db	454	rkmllnpnkdplvlgvrvwvkaipqflvghldlldsakmalrdsqfnglflgnyvsvv	513		

OY 523 ALGRCEGAYEVAEFTGELSRAYK 548
 DB 514 algrcegayevaevkeflsagyak 539

RESULT 6
 AAB72907 standard; Protein: 539 AA.

ID AAB72907 standard; Protein: 539 AA.
 AC AAB72907;
 DT 16-MAY-2001 (first entry)

DE Cotton protoporphyrinogen oxidase SEQ ID NO: 16.
 DE Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
 KM soybean; sugar beet; oilseed rape; sugar cane; mutant; mutain.
 XX Gossypium hirsutum.
 OS
 XX WO200112825-A1.
 PN
 XX 22-FEB-2001.
 PD
 XX 30-JUN-2000; 2000WO-EP06127.
 PF
 XX 13-AUG-1999; 99US-0373691.
 PR
 XX (SYNG-) SYNGENTA PARTICIPATIONS AG.
 PA
 XX Johnson MA, Volrath SL, Helfetz PB, Law MD;

PI WPI: 2001-234914/24.
 DR N-PSDB; AAF76579.
 DR
 XX Plant DNA molecules encoding herbicide-tolerant forms of
 PT protoporphyrinogen oxidase which are useful for rationally designing
 FT new inhibitory herbicides and for producing herbicide-tolerant
 FT transgenic plants and seeds
 XX Claim 22; Page 184-187; 228pp; English.

XX The present invention provides the protein and coding sequences of a
 CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
 CC sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
 CC oxidase (protox) enzyme. Examples of these mutants are shown in
 CC AAB72920-AAB72926. They are useful as they enable the production of
 CC herbicide-tolerant plants and seeds. The present sequence is a protox
 CC protein.
 XX
 PS Sequence 539 AA;

Query Match 74.3% Score 2113; DB 22; Length 539;
 Best Local Similarity 79.2%; Pred. No. 2.8e-194;
 Matches 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;

OY 47 WRRCVAKDYVPPSAVDGGPAALDQVYVAGISGLCIAQVMSANV---PMLMATEA 102
 DB 34 FKIRCslaegpLtsSSkIdggssIadcvIv99gIsgIcaqalatkhrdvaahvIylea 93
 OY 103 RDRAGNITTVEDGGLWMEGPNFOPSDPMLTMVAVDCGLKDDLVLDGPNAPRFVLMKGG 162
 DB 94 rdrvggnlttverdyglweegpnstfqsdpdlImavdsqIkdldlylgdnprrfvlwegk 153
 OY 163 LRVPKRLDLPFPPIMSTPKLRAGPRTGIRSPSPGHESEVDFVRNMGGEVFEPLI 222
 DB 154 lrvpksrldlpfdlmsrtaqkIraagIgaIgrIpppyeeevetvrmIgaevIefI 213
 OY 223 EPRFCGVVYVGDSDSKLSMKARAGKVKWKLERTGSGTIGGTEFKAIKRSSTPKAPRDRPKP 282
 DB 214 eprfcgvvaygdpsklsmkaraIgrvkwkIleI99sIlggtfklrgermktkppkprdrIpkp 273

OY 283 KQGVGSEFRKGLRMLPDLISARLGSKLSKLSITSEKGVHLYEFTPEGVVSLQSR 342
 DB 274 kgvysfrkglrmlpdlisarlgsksksksitsekghlyeftepegvvs1qsr 333
 OY 343 SIYMWPSYVANSNIIIRPLSVAADALSNFYPPVGAVTITSPQEARIDRERLVJGELKFG 402
 DB 334 svmtlpsvsnllhplsaadalsqfyppvasvsvypkaeirkeclldgkfg 393
 OY 403 QLRPRGCVETLCTIYSSSFPNRAKGRVLLIYTGAKKPELISKEQVLEVVNDRL 462
 DB 394 qlrprsglctlytssslfprnarkgrvlllytgakkpeliskEQVleevndrl 453
 OY 463 RKMILKPKAQPDLVGVVWVWVQAIPOFLVGHLDPLSTAKAAMNDNGLEGLFGSNVYSGV 522
 DB 454 rkmilpnaqkdpIvgyvwpkaiPqflvghld1ldsakmalrdsqfIflg9nyvsgv 513
 OY 523 ALGRCEGAYEVAEFTGELSRAYK 548
 DB 514 algrcegayevaevkeflsagyak 539

RESULT 7
 AAB72921 standard; Protein: 539 AA.

ID AAB72921 standard; Protein: 539 AA.
 AC AAB72921;
 DT 16-MAY-2001 (first entry)

DE Cotton protoporphyrinogen oxidase Y428H mutant.
 DE Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
 KM soybean; sugar beet; oilseed rape; sugar cane; mutant; mutain.
 XX Gossypium hirsutum.
 OS
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Msc-difference 428
 FT /note="Wild-type Tyr substituted by His"
 XX
 XX WO200112825-A1.
 PN
 XX 22-FEB-2001.
 PD
 XX 30-JUN-2000; 2000WO-EP06127.
 PF
 XX 13-AUG-1999; 99US-0373691.
 PR
 XX (SYNG-) SYNGENTA PARTICIPATIONS AG.
 PA
 XX Johnson MA, Volrath SL, Helfetz PB, Law MD;

PI WPI: 2001-234914/24.
 DR
 XX Plant DNA molecules encoding herbicide-tolerant forms of
 PT protoporphyrinogen oxidase which are useful for rationally designing
 FT new inhibitory herbicides and for producing herbicide-tolerant
 FT transgenic plants and seeds
 XX Claim 22; Page 1; 228pp; English.

XX The present invention provides the protein and coding sequences of a
 CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
 CC sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
 CC oxidase (protox) enzyme. Examples of these mutants are shown in
 CC AAB72920-AAB72926. They are useful as they enable the production of
 CC herbicide-tolerant plants and seeds. The present sequence is a mutant
 CC protein.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from that given in SEQ ID NO: 16 (see AAB72907).
 CC
 PS Sequence 539 AA;

Query Match 74.1% Score 2108 DB 22 Length 539
Best Local Similarity 79.1% Pred. No. 8 5e-194
Matches 400; Conservative 47; Mismatches 55; Indels 4; Gaps 1;

QY	47	WRTRGVAKDITYVPSANVGDGPAALDCCYVYAGISGLCICIAQVMSANV----	PNLKWTGA	102
DB	34	fkfrcslaeqrltsskldgsssladcvlyvgglsjclqalatkhrdvasnvlvtea	93	
QY	103	RDRAGNTTVERHGDYLMWEGEENSRFPSPDMLTMAVDCCGLKDDVYLDGPFNAFRRVLMWKGK	162	
DB	94	rdrvggnltvverdyllweepnstfcpdpdlmtaavsglkddlvlygdnpvrfvlywggk	153	
QY	163	LRPVPSKLTLDLPEFDLMSIPGKLRAFGPPIGLRPSPPGHEESVQGFVRRNLGGVEFERLI	222	
DB	154	lrpvpkpldldpfdlmslqgklraqfgalqirpppyeesevfevtrnliqaevfefli	213	
QY	223	EPCSGYVYGDPSKLSMKAAFGVWKLKEMGGSTIGTFRKAIKSSSPPKARPDRPKP	282	
DB	214	epfcsyvyagdpkslsmkaafgyrvwkleegslvgstlqgtrkcpkpprpprlpkp	273	
QY	283	KGQTVGSFRRKRLMPLDAISARLSGSKLSMKLUSSTTRSEKGGYHLTYERPBGVVSLSQR	342	
DB	274	kgqtvsfrkqrlmlpaalansgnvklswklsstlknggyltletpbgmvsiqsr	333	
QY	343	SIWTVPSYVANSILRPLSVAAADALSNFYYPVGAVNTISYPOEAIQDRLVDGELKGFQ	402	
DB	334	svwmltprshvasnllhplsaaadalsqfyfypvavsvtvsypkealtrkcollidgslkgfg	393	
QY	403	QLHPRTOGVERLTGTYSSSLFPNRAKGRVILLNLTIGGAKNPEILISKTESQVVEVDRDL	462	
DB	394	qllhprsgletqlytlyssslfpnrapsgrvllnhhlgagatngslsktegelveavdrdl	453	
QY	463	RKMLIKPKADPELWVGRVWVQAIPOELVGHLDLTSRAKAAMNDGLEPLGNGVYSGV	522	
DB	454	rkmllmpnkapdlvlygrvwpkaiqqilvghldllsaskmalrdsqfhljlygnyvsgv	513	
QY	523	ALGRCVEGAYEVAEVTGFLSRAYK	548	
DB	514	algrcvegayeavaevkflisqayk	539	

RESULT 8
AAB72922 standard: Protein: 539 AA.
AAB72922:
AAB72922:
16-MAY-2001 (first entry)
Cotton protoporphyrinogen oxidase Y428a mutant.
Protoporphyrinogen oxidase: protox; herbicide-tolerance; wheat; rice;
soybean; sugar beet; oilseed rape; sugar cane; mutant; mutlein.
Gossypium hirsutum.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 428 /note= "wild-type Tyr substituted by Ala"
PN WC200112825-A1.
PD 22-FEB-2001.
XX 30-JUN-2000: 2000MO-EP06127.
XX 13-AUG-1999: 99US-0373691.
XX (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX PA

Query Match 74.0% Score 2104 DB 22 Length 539
Best Local Similarity 79.1% Pred. No. 2 1e-193
Matches 400; Conservative 46; Mismatches 56; Indels 4; Gaps 1;

PI	Johnson MA, Volrath SL, Helfetz PB, Law MD: WPI: 2001-234914/24.			
XX				
DR				
XX				
PT	Plant DNA molecules encoding herbicide-tolerant forms of protoporphyrinogen oxidase which are useful for rationally designing new inhibitory herbicides and for producing herbicide-tolerant transgenic plants and seeds -			
PT	Claim 22; Page -: 228pp; English.			
XX				
PS				
CC	The present invention provides the protein and coding sequences of a number of herbicide-tolerant forms of wheat, soybean, rice, sorghum, sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen oxidase (protox) enzyme. Examples of these mutants are shown in AAB72920-AAB72926. They are useful as they enable the production of herbicide-tolerant plants and seeds. The present sequence is a mutant protox protein.			
CC	Note: The present sequence is not shown in the specification but is derived from that given in SEQ ID NO: 16 (see AAB72907).			
CC				
XX				
XX				
SQ	Sequence 539 AA:			
QY	47	WRTRGVAKDITYVPSANVGDGPAALDCCYVYAGISGLCICIAQVMSANV----	PNLKWTGA	102
DB	34	fkfrcslaeqrltsskldgsssladcvlyvgglsjclqalatkhrdvasnvlvtea	93	
QY	103	RDRAGNTTVERHGDYLMWEGEENSRFPSPDMLTMAVDCCGLKDDVYLDGPFNAFRRVLMWKGK	162	
DB	94	rdrvggnltvverdyllweepnstfcpdpdlmtaavsglkddlvlygdnpvrfvlywggk	153	
QY	163	LRPVPSKLTLDLPEFDLMSIPGKLRAFGPPIGLRPSPPGHEESVQGFVRRNLGGVEFERLI	222	
DB	154	lrpvpkpldldpfdlmslqgklraqfgalqirpppyeesevfevtrnliqaevfefli	213	
QY	223	EPCSGYVYGDPSKLSMKAAFGVWKLKEMGGSTIGTFRKAIKSSSPPKARPDRPKP	282	
DB	214	epfcsyvyagdpkslsmkaafgyrvwkleegslvgstlqgtrkcpkpprpprlpkp	273	
QY	283	KGQTVGSFRRKRLMPLDAISARLSGSKLSMKLUSSTTRSEKGGYHLTYERPBGVVSLSQR	342	
DB	274	kgqtvsfrkqrlmlpaalansgnvklswklsstlknggyltletpbgmvsiqsr	333	
QY	343	SIWTVPSYVANSILRPLSVAAADALSNFYYPVGAVNTISYPOEAIQDRLVDGELKGFQ	402	
DB	334	svwmltprshvasnllhplsaaadalsqfyfypvavsvtvsypkealtrkcollidgslkgfg	393	
QY	403	QLHPRTOGVERLTGTYSSSLFPNRAKGRVILLNLTIGGAKNPEILISKTESQVVEVDRDL	462	
DB	394	qllhprsgletqlytlyssslfpnrapsgrvllnhhlgagatngslsktegelveavdrdl	453	
QY	463	RKMLIKPKADPELWVGRVWVQAIPOELVGHLDLTSRAKAAMNDGLEPLGNGVYSGV	522	
DB	454	rkmllmpnkapdlvlygrvwpkaiqqilvghldllsaskmalrdsqfhljlygnyvsgv	513	
QY	523	ALGRCVEGAYEVAEVTGFLSRAYK	548	
DB	514	algrcvegayeavaevkflisqayk	539	

RESULT 9
AAB72923 standard: Protein: 537 AA.
AAB72923:
AAB72923:
16-MAY-2001 (first entry)
XX
DT
XX

DE Arabidopsis protoporphyrinogen oxidase A2201 mutant.

XX XX Arabidopsis protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;

KM soybean; sugar beet; oilseed rape; sugar cane; mutant; mutlein.

XX Arabidopsis thaliana.

OS Synthetic.

XX XX

FH Key Location/Qualifiers

FT Misc-difference 220 /note="wild-type Ala substituted by Ile"

XX MO200112825-A1.

XX PD 22-FEB-2001.

XX PF 30-JUN-2000; 2000WO-EP06127.

XX PR 13-AUG-1999; 99US-0373691.

XX PA (SYNGENTA PARTICIPATIONS AG.

XX PI Johnson MA, Volrath SL, Helfetz PB, Law MD;

XX DR WPI: 2001-234914/24.

XX PT Plant DNA molecules encoding herbicide-tolerant forms of protoporphyrinogen oxidase which are useful for rational design of new inhibitory herbicides and for producing herbicide-tolerant transgenic plants and seeds

XX PS Claim 23; Page -: 228pp; English.

CC The present invention provides the protein and coding sequences of a number of herbicide-tolerant forms of wheat, soybean, rice, sorghum, sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen oxidase (protoporphyrinogen oxidase) enzyme. Examples of these mutants are shown in AAB72920/AAB72926. They are useful as they enable the production of herbicide-tolerant plants and seeds. The present sequence is a mutant proox protein.

CC Note: The present sequence is not shown in the specification but is derived from that given in SEQ ID NO: 2 (see AAB72901).

XX CC

XX SQ Sequence 537 AA;

Query Match 73.1%; Score 2078; DB 22; Length 537;

Best Local Similarity 75.4%; Pred. No. 6.5e-191;

Matches 404; Conservative 50; Mismatches 72; Indels 10; Gaps 4;

OY 22 IAFLNRTSFTPFSSISRSNVCNGW---RTRCSYAKDYTPVSSAVDGGPAEL--DCVI 76

DB 3 ILLIPLTQSLIPLSFSKPN-ILINVKPILRCVAGGPRVGVGSSKIEG99GTTTITDDEV1 61

OY 77 VGAGISGLCIAQVMSANY----DNLMWTEARDRAGNITTVTERDGYLMBEGRNSQPSPDP 132

DB 62 VGGVGIISGLCIQATGATKHPDAENLIVTEAKKRVGNLITTEENGFLIWEEGPSTFQSPDP 121

OY 133 MITMNVDCGKIDDLVLGDDPNARREVWKKLRVPSKLTDPDPLDMISFGKIRAFQEI 192

DB 122 MTLNVVDSGILKXDDLVDLQPRAPRITVLW9KILRPVPSKLTDLPIFLDMS19GKILTAG1 181

OY 193 GLRPSRPGHEBSVBOFVRRNLGGEVFERLIEPFGCSGVYVGDPSKLSKKAFAFKWKLLEBT 252

DB 182 GILRPDPREEVVEFVTRNLIDGVFERRLIEPFCGGVYIGDPKLSMKKAFAFGKVLWVLEGN 241

OY 253 GSSIIIGGFFAKIKERSSTPKARDPRIKPKGGQTVGSPFRKIGIRMLDPAIAGARSKLKIKLS 312

DB 242 GSSIIIGGFFAKIKERSSTPKARDPRIKPKGGQTVGSPFRKIGIRMLDPAIAGARSKLKIKLS 301

OY 313 WKLSSITKSEKGGVNLLETPREBVGVSIDSRSITVMVPRVYVSNLIRELVVAADALSNRY 372

DB 302 WKLSSITKSEKGGVNLLETPREBVGVSIDSRSITVMVPRVYVSNLIRELVVAADALSNRY 361

OY 373 YPPVGAVTITSPQDAIRDRERYDGLKGPQGLHRRPQQGVETLCTIYSSLSLFPNRARPKGRV 432

DB 362 YPPVGAVTSIPXEAIRLTELIDGELK9IGLHPRLGGVETLGLYSSSLFPNRPDPRI 421

OY 433 ILLNVIYGAKNRPETLSTKTESOLEVVDRLRKMILIKPPADDPVSGVRWPPQAIPOPLVG 492

DB 422 ILLNVIYGSTNTGLIISKEGELVEAVDRLIKKPNSTDPKLGVRWPPQALDQFLIVG 481

OY 493 HUDPSTAKAAMNDGLEGPEFLGQNVYSVALGRCVGAEEVSEVTFGLSHYAVK 548

DB 482 HFDLIDTAKESLSSGYEGLIIGNYVAGVALGRCVGAEEVSEVTFGLSHYAVK 537

RESUltm 10

AAR90295

ID AAR90295 standard; Protein; 537 AA.

XX AAR90295;

AC AAR90295;

XX 02-JUL-1996 (first entry)

DE Protein having protoporphyrinogen oxidase activity.

XX KM

KM Protoporphyrinogen oxidase; PPO; herbicide; resistance; mutant

KW variegate porphyria.

XX OS Arabidopsis sp.

XX PN W09534659-A1.

XX PD 21-DEC-1995.

XX PE 08-JUN-1995; 95MO-IB00452.

XX PR 16-JUN-1994; 94US-0261198.

XX PA (CIBA) CIBA GEIGY AG.

XX PI Volrath S, Ward ER;

XX DR WPI: 1996-049687/05.

DR N-PSDB; AAT11673.

XX DNA encoding eukaryotic proto:porphyrinogen oxidase and herbicide resistant mutants - used to make herbicide resistant plants and for diagnosis and treatment of variegate porphyria

PT

XX Claim 6; Page 71-73; 118pp; English.

CC DNA encoding a modified protoporphyrinogen oxidase (PPO) or a chimeric gene comprising a promoter (pref. active in a plant) linked to such a DNP or DNA encoding a wild type PPO can be used (1) to impart herbicide resistance to plants; (2) for treating and CC diagnosing deficient PPO activity in animals (esp. variegate porphyria); and (3) for the production of recombinant PPO which is useful as an assay reagent and in rational design of new inhibitory CC herbicides. Herbicide resistant PPO genes can also be used to CC select plants transformed with a transgene and probes derived from the genes can be used to quantify levels of PPO mRNA.

XX CC

XX SQ Sequence 537 AA;

Query Match 73.0%; Score 2075; DB 17; Length 537;

Best Local Similarity 75.4%; Pred. No. 1.3e-190;

Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

OY 22 IAFLNRTSFTPFSSISRSNVCNGW---RTRCSYAKDYTPVSSAVDGGPAEL--DCVI 76

DB 3 ILLIPLTQSLIPLSFSKPN-ILINVKPILRCVAGGPRVGVGSSKIEG99GTTTITDDEV1 61

OY 77 VGAGISGLCIAQVMSANY----DNLMWTEARDRAGNITTVTERDGYLMBEGRNSQPSPDP 132

DB 62 VV9GISGLCIAGALAKHPDAAPNLIYVEAKDRYGNALITREENGFLWEQGNSTFQPSDP 121
 OY 133 MLTMAVDCGLKDDLVLDGPNAPRFVLMKGLRPVPSKLTDLPEFDLMSIPKLRAGFCGPI 192
 DB 122 mltnvvdsgjkdldvlvgdptlprfvlmngkrlrpkpqqgvtgsvftrkylmpealsarlsqkvlis 301
 OY 193 GLRPPSGHESEVDFVRRNLDGEGVFERLIEPFCGQVYVGPFSKLSMKAARFKWVAKLEET 252
 DB 182 glrppsgheesveefvrrnldgdevferlliepcsgvyvagdpskslsmkaadgkvwklegn 241
 OY 253 GGSIIIGTFFKAIKERSPPKAPRDRPLPKPKGQTVGSFRRKGLMLPDAISARLSKLUKLS 312
 DB 242 ggsiilggtfkaiqerknapkkaerdrprkpkqqgvtgsvftrkylmpealsarlsqkvlis 301
 OY 313 WKLSSTIRSEKGGYHLTYETPEGVVSLQSRISYVTVVPSYVANSILRPLSVAAADALSIFY 372
 DB 302 wklsstirsekgyhltvetpegvvsllqsrivtvvpsvyvansilrplsvaaadalsify 361
 OY 373 YPPVGAVTISYPOEAIKRDRLVLDGELKRGQQLHPRTOGVTGTTGTYSSSLFPNRAKGRV 432
 DB 362 yppvaavsisypkealrreclldgelkrgfqlhprtcqgvtclgclyssslfpnrappgrv 421
 OY 433 LLNMTYIGAKNPEILSKTESQDLVEVVDRLRKMILKPKAODPLVGVVWVWVQALPQFLV 492
 DB 422 llmnylgsntglilsksegevlveavdrdrkmlkpnstcdplklygvrvwvwpqalpqflvg 481
 OY 493 HLDLSTAKAAMNDNGLEGLFVGNVYSGVALGRVCEGAYEVAEVTGFLSRAYK 548
 DB 482 hfdldtaksstlssgyeglfignyvayvalgrvcvegayetalavnmfmryayk 537

RESULT 11
 AAM41603 standard; Protein: 537 AA.
 ID AAM41603;
 AC AAM41603;
 DI 20-APR-1998 (first entry)
 DE Arabidopsis thaliana protox-1.
 XX
 KW Protoporphyrirogen oxidase-1; protox-1; promoter;
 KW herbicide resistance; breeding programme; probe; gene isolation;
 KW genomic mapping.
 OS Arabidopsis thaliana.
 XX
 PN W09732028-A1.
 PD 04-SEP-1997.
 PE 27-FEB-1997; 97WO-US03343.
 PR 21-JUN-1996; 96US-0020003.
 PR 28-FEB-1996; 96US-0012705.
 PR 28-FEB-1996; 96US-0013612.
 PA (NOVS) NOVARTIS AG.
 PI Johnson MA, Volrath SL, Ward ER;
 DR WPI: 1997-489209/45.
 DR N-PSDB; AAV04304.
 XX
 FT DNA containing a plant proto-porphyrinogen oxidase gene promoter -
 FT optionally linked to a heterologous gene, especially to express
 FT herbicide-resistant enzymes, and plants containing such constructs
 XX
 PS Claim 27; pages 36-39; 11app; English.
 CC The present sequence is Arabidopsis thaliana
 CC protoporphyrirogen oxidase-1 (protox-1).

CC The protox-1 promoter can be used to express herbicide resistant
 CC enzymes, specifically protox, i.e. a plant tissue, plant or progeny
 CC containing a chimeric gene of the promoter and a heterologous
 CC coding sequence. The plant can also be used in breeding programmes.
 CC Also hybridising fragments of the protox coding sequence can be
 CC used as probes, e.g. to isolate related genes or for genomic
 CC mapping.
 XX
 SQ Sequence 537 AA:
 OY 22 LABLNRTSIFPSSISKRNSVNCNGW---RTRCSYAKDYTVSSAVDGGPAALR--DCVI 76
 DB 3 lslrlptqslilpfskgn-lrlnvylkrlrcsvagpvtvsgskiegqggtlttdcyl 61
 OY 77 VGAGISGICIAQVMSANY---PMLWTEARDRAGGNTITVYRDSGLYMBEGRPNFQPSDP 132
 DB 62 vv9gisglciagalakhpdaapnliyteakdrvgnalltreengflweqgnstfqsdp 121
 OY 133 MLTMAVDCGLKDDLVLDGPNAPRFVLMKGLRPVPSKLTDLPEFDLMSIPKLRAGFCGPI 192
 DB 122 mltnvvdsgjkdldvlvgdptlprfvlmngkrlrpkpqqgvtgsvftrkylmpealsarlsqkvlis 301
 OY 193 GLRPPSGHESEVDFVRRNLDGEGVFERLIEPFCGQVYVGPFSKLSMKAARFKWVAKLEET 252
 DB 182 glrppsgheesveefvrrnldgdevferlliepcsgvyvagdpskslsmkaadgkvwklegn 241
 OY 253 GGSIIIGTFFKAIKERSPPKAPRDRPLPKPKGQTVGSFRRKGLMLPDAISARLSKLUKLS 312
 DB 242 ggsiilggtfkaiqerknapkkaerdrprkpkqqgvtgsvftrkylmpealsarlsqkvlis 301
 OY 313 WKLSSTIRSEKGGYHLTYETPEGVVSLQSRISYVTVVPSYVANSILRPLSVAAADALSIFY 372
 DB 302 wklsstirsekgyhltvetpegvvsllqsrivtvvpsvyvansilrplsvaaadalsify 361
 OY 373 YPPVGAVTISYPOEAIKRDRLVLDGELKRGQQLHPRTOGVTGTTGTYSSSLFPNRAKGRV 432
 DB 362 yppvaavsisypkealrreclldgelkrgfqlhprtcqgvtclgclyssslfpnrappgrv 421
 OY 433 LLNMTYIGAKNPEILSKTESQDLVEVVDRLRKMILKPKAODPLVGVVWVWVQALPQFLV 492
 DB 422 llmnylgsntglilsksegevlveavdrdrkmlkpnstcdplklygvrvwvwpqalpqflvg 481
 OY 493 HLDLSTAKAAMNDNGLEGLFVGNVYSGVALGRVCEGAYEVAEVTGFLSRAYK 548
 DB 482 hfdldtaksstlssgyeglfignyvayvalgrvcvegayetalavnmfmryayk 537

RESULT 12
 AAM25746
 ID AAM25746 standard; Protein: 537 AA.
 AC AAM25746;
 DI 01-MAR-1998 (first entry)
 DE Arabidopsis protoporphyrirogen oxidase (protox-1).
 XX
 KW Protox-1; protoporphyrirogen oxidase; inhibitor;
 KW herbicide tolerance; herbicide resistance; transgenic plant.
 OS Arabidopsis thaliana.
 XX
 PN W09732011-A1.
 PD 04-SEP-1997.
 PE 27-FEB-1997; 97WO-US03313.
 XX

PR 21-JUN-1996; 96US-0020003.
 PR 28-FEB-1996; 96US-0012705.
 PR 28-FEB-1996; 96US-0013612.
 XX
 XX (NOVS) NOVARTIS AG.
 XX
 XX Helferz PB, Johnson MA, Potter SL, Volrath SL, Ward ER;
 PI WPI: 1997-448683/41.
 DR N-PSDB: AAT86129.
 XX
 XX New DNA encoding plant protoporphyrinogen oxidase enzyme - and
 PI herbicide resistant mutants, useful to prepare plants resistant to
 PI herbicide which therefore kills undesired vegetation only
 XX
 PS Claim 95; Page 95-98; 196pp; English.
 CC This protein comprises Arabidopsis protoporphyrinogen oxidase
 CC (protox-1), an enzyme that catalyses the oxidation of
 CC protoporphyrinogen IX to protoporphyrin IX. Its amino acid sequence
 CC was deduced from a previously obtained Protox-1 cDNA clone (see
 CC AA896129). Arabidopsis protox-1 can be modified to render it
 CC resistant to protox inhibitors and hence tolerant of certain
 CC herbicides. Plants, especially crop plants, may be engineered for
 CC resistance to protox inhibitors via mutation of the native protox
 CC gene to a resistant form, or they may be transformed with a gene
 CC encoding an inhibitor-resistant form of a plant protox enzyme, such
 CC as claimed forms from wheat, soybean, cotton, sugarbeet, oilseed
 CC rape, rice and sorghum (see AW25738-48). Application of herbicide
 CC will then kill undesired vegetation only. Protox enzymes can also
 CC be expressed in transformed host cells and used to identify
 CC inhibitors of protox enzyme activity, i.e. herbicide candidates, or
 CC to design herbicide tolerant forms of the enzyme.
 CC
 SO Sequence 537 AA:
 Query Match 73.0%; Score 2075; DB 18; Length 537;
 Best Local Similarity 75.4%; Pred. No. 1.3e-190;
 Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

QY 493 HLDPTLSTAKAAANDNGLEFLGGRVYSGVALGRCEGAYEVAESEYTGELSRAYK 548
 Db 482 hfdldtksksslsssgyegflfignyvavagvalgrcvegayetalevnmimsryayk 537
 RESULT 13
 ID AAW51347 standard; Protein; 537 AA.
 AC AAW51347;
 XX
 XX 14-AUG-1998 (first entry)
 DE Arabidopsis proto-porphyrinogen oxidase-1 (protox-1).
 XX
 XX Proto-porphyrinogen oxidase; protox; herbicide tolerant; resistance;
 KW inhibitor; variegated porphyrin.
 OS Arabidopsis thaliana.
 XX
 XX US5767373-A.
 PD 16-JUN-1998.
 XX
 XX 16-JUN-1994; 94US-0261198.
 PF 16-JUN-1994; 94US-0261198.
 PR 06-JUN-1995; 95US-0472028.
 PR 16-JUN-1994; 94US-0261198.
 XX
 PA (NOVS) NOVARTIS FINANCE CORP.
 PI Volrath S, Ward ER;
 DR WPI: 1998-361821/31.
 DR N-PSDB: AAV07251.
 XX
 XX DNA encoding eukaryote herbicide resistant proto-porphyrinogen
 PT oxidase - useful for producing recombinant plants having functional
 PI enzyme, to be grown in the presence of herbicides
 PS Claim 6; Columns 21-22; 43pp; English.
 XX The invention relates to eukaryotic DNA sequences coding for native
 CC proto-porphyrinogen oxidase (protox) or modified forms of the enzyme
 CC which are herbicide tolerant. Plants having altered protox activity
 CC which confers tolerance to herbicides are also provided. These plants
 CC may be bred or engineered for resistance to protox inhibitors via
 CC mutation of the native protox gene to a resistant form or through
 CC increased levels of expression of the native protox gene, or they may be
 CC transformed with modified eukaryotic or prokaryotic protox coding
 CC sequences or wild type prokaryotic protox sequences which are herbicide
 CC tolerant. In the human condition variegated porphyrin, an autosomal
 CC dominant disorder characterised by neuropsychiatric symptoms, decreased
 CC protox activity is detected. Thus the DNA molecules can be used to
 CC provide probes to detect and quantify protox levels in the diagnosis of
 CC diseases associated with the enzyme. The present sequence represents
 CC Arabidopsis proto-porphyrinogen oxidase-1 (protox-1).
 CC
 SO Sequence 537 AA:
 Query Match 73.0%; Score 2075; DB 19; Length 537;
 Best Local Similarity 75.4%; Pred. No. 1.3e-190;
 Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

PF 30-JUN-2000; 2000MO-EP06127.
 XX
 PR 13-AUG-1999; 99US-0373691.
 XX
 PA (SYNG-) SYNGENRA PARTICIPATIONS AG.
 XX
 PI Johnson MA, Volrath SL, Helfetz PB, Law MD;
 XX
 DR WPI: 2001-234914/24.
 DR N-PSDB; AAF76571.
 XX
 XX Plant DNA molecules encoding herbicide-tolerant forms of
 PT protoporphyrinogen oxidase which are useful for rationally designing
 PT new inhibitory herbicides and for producing herbicide-tolerant
 PT transgenic plants and seeds -
 XX
 PS Claim 23; Page 137-140; 228pp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
 CC sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
 CC oxidase (protox) enzyme. Examples of these mutants are shown in
 CC AAB72920-AAB72926. They are useful as they enable the production of
 CC herbicide-tolerant plants and seeds. The present sequence is a protox
 CC protein.
 CC
 XX
 SQ Sequence 537 AA;

Query Match 73.0%; Score 2075; DB 22; Length 537;
 Best Local Similarity 75.4%; Pred. No. 1.3e-190;
 Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

QY 22 LAPLNRTSFIFPFSSISKRNSVNCNGW---RTRCSVAKDYTVSSSAVDGGPAEL--DCVY 76
 Db 3 Isllrplrtgllpsfiskpn-lrlnvkplrllrcsvag99plrvsskskleg99gttlctdcv1 61
 QY 77 VGAGISGLCIAQVMSANY---PNIWVTEARDRAGNITTVERRDGYLMEEGPNSFQPSPDP 132
 Db 62 vgg9igs9lciagatatkhpdaapnllvteaktvgnlitrreengflweegpnsfqpsdp 121
 QY 133 MLTMAPDCCGLKDDLVLGDPNAPRPVLMKGLRPVPSKLRDLPFPDLMSIPGKLRAGFGPT 192
 Db 122 mltnvvvsg1kddlvlgdpaprfrlwn9k1rvvpsk1tdlppfdlms1gk1ragf9al 181
 QY 193 GLRPSPEGHESYEOAFRRLVGGVFERLIEPCSGVYVGDPSKLSMKAAFGKWKLEEF 252
 Db 182 glp9pp9resvveefvrrnlqdevterliepfcsvyagdp9k1smkka9fgkvwk1eqn 241
 QY 253 GGSIIIGGTFRAIKERSSTPKAPRPDPLPKPKGQTGVSFRKGLRMLPDATISARLGSKILKS 312
 Db 242 g9s9l9gtfkaiqrknapkaerdp9r9k9g9gtvgsfrkg1rmlpeaisar1gskvk1s 301
 QY 313 WKSISTRSKSGYHLYTTEPEGVVSLQSRSTVMTVPSTYVANSIIIRPLSYAAADALSNEY 372
 Db 302 wk1sg9l9tkles99ynll9ep9d9lvs9vsksvmtvps9hsa99ll9rpl9sesaana1sk1y 361
 QY 373 YPPYGAVTISYPOEATRDERLVDGELKFGQLHPTQGVETLGTITVSSSLFPPRARPKGRY 432
 Db 362 yppvaav9l9y9kaaitelcl9gel9kg9dl9p9rc9vecl9gl9ys9sl9f9n9rap9gl 421
 QY 433 LLNNTYGGAANNPRLSKTESQLVAVVDRDLKMLIKPKAODPLVGVGRVWPQAIPQFLVNG 492
 Db 422 lllnyl9gstnt9l9sk9egel9veavdrdl9rknll9kpn9t9dp19k19vrvwp9a19p9fl9v 481
 QY 493 HLDLTSTAKAAMNDNGIEGLFLGNTVSGVALGRCVGADEVASEVTGFSRTRAKK 548
 Db 482 htfdl9tdak9s9l9ts9sgyegfl9gn9y9va9l9rc9ve9a9el9vna9lm9m9s9y9ak 537

Search completed: July 3, 2001, 10:32:56
 Job time: 1975 sec

OM of: US-09-508-418-2 to: GenBank: * out_format : pfs
Date: Jul 3, 2001 11:36 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-DEB=xlp
-DB=genbank
-MINARCH=0.100
-IOOPT=0.000
-IOCPRT=0.000
-IOGAP=4.500
-IOGAPEXT=0.050
-XGAPOP=10.000
-XGAPEXT=0.500
-FGAPOP=6.000
-FGAPEXT=7.000
-YGAPOP=10.000
-YGAPEXT=0.500
-DELDP=6.000
-DELEX=7.000
-START=1
-MATRIX=blonsum2
-TRANS=humana4.cdi
-LIST=45
-DOCALIGN=200
-THR_SCORE=pcr
-THR_MAX=100
-THR_MIN=0
-ALIGN=15
-DOCAL=LOCAL
-OUTFRM=pfs
-NORM=ext
-MINLEN=0
-MAXLEN=2000000000
-USER=us09508418
-GENI_1_5606
-NCPU=6
-ICPU=3
-LONGLOG
-NO_XLPRX
-WAIT
-THREADS=1

Search Information block:
Query: US-09-508-418-2
Database: GenBank: *
Query Length: 548

Database sequences: 1344157
Database Length: 856060004
Search time (sec): 1837.380000

Table with columns: Sequence, Strd Orig, ZScore, EScore, Len, Documentation. Contains sequence alignment data for various Nicotiana and Solanum species.

Table with columns: seq_name, seq_documentation_block, LOCUS, DEFINITION, ACCESSION, VERSION, KEYWORDS, ORGANISM, REFERENCE, AUTHORS, TITLE, JOURNAL, FEATURES, source. Contains detailed information for sequence gb_D11:AF044128.

seq_name: gb_D11:AF044128
seq_documentation_block:
LOCUS AF044128 1874 bp mRNA
DEFINITION Nicotiana tabacum protoporphyrinogen oxidase PX-1 mRNA, complete cds.
ACCESSION AF044128
VERSION AF044128.1 GI:4105185
KEYWORDS
common tobacco.
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; euasterids I; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 1874)
Horikoshi,M., Mametsuka,K. and Hirooka,T.
The Molecular Basis of Photoleaching Herbicide Resistance in tobacco
Unpublished
2 (bases 1 to 1874)
Horikoshi,M., Mametsuka,K. and Hirooka,T.
Direct Submission
Submitted (22-JAN-1998) Research Center, Nihon Nohyaku Co. Ltd., 345, Oyamada-cho, Kawachi-nagano, Osaka 586-0094, Japan
Location/Qualifiers
1..1874
/organism="Nicotiana tabacum"
/cultivar="SRI"
/db_xref="taxon:4097"
26..1672
/EC_number="1.3.3.4"
/codon_start=1
/product="protoporphyrinogen oxidase px-1"
/protein_id="AAD02290.1"
/db_xref="GI:4105186"
/translaton="MTRPPIAHNPIFTHQSSSSPLAFLNRTSFTFPSSISKRNSVNC
NGRTRCSVAKDYTVPSAVDGGPAPALDVCVIVAGISGLCAQVMSAVYPLMVTGA
RDPAGGNITTVRDYIMRGGVEMRPPSPRPPSPALMTMVDGLKDIVGDNARFVLMK
GKRIPVPSKTTIDIPFDLMSIRPKLRAGGAIAGLRSPPHGSESVQFRRMLGGEV
ERLLEPKSGVAVGDPKSLMSKAAFGKWKLEDTGSSITGGFKAKESSTPKAPRD
PRLEPKRQGVGFRRKRLMPLPAISARVLSKIKLSSTTKRSKGGYHLVTEPE
GVVSLQSRISVIMTVPSYVSNILIRPLVSAVAANLNFYPPGAVTISYPOAIBDER
LVNGELKRGFGLHPPTGVEYETLGTIVSSIFPRAKGGVLLNLYTGAKNEELISK
ESQLVFVYVDRLEKMLIKPKAODPIVGVVMPQALDPPDLYVHDDTLSTRAKAMNDG
LEGLFAGNYSVVALRGVRYEVASEVYTFLSRYAK"

alignment_scores:
Quality: 2832.00 Length: 548
Ratio: 5.177 Gaps: 0
Percent Similarity: 99.818 Percent Identity: 99.635

alignment_block:
US-09-508-418-2 x AF044128 ..
Align seg 1/1 to: AF044128 From: 1 to: 1874
1 Metrrrrrrrthprrllalashnhsprknsnllpethrhhisglase 17
|||||
26 ATGAAACACACACCCATGCGCCATCATATATTTTTCACCTCACCAAGTC 75
|||||
17 rrserserserproleualaphelvausnargtthserphelilepropes 34
|||||
76 GTCGTCATCGCCATTCGATTCCTTAACCGAAGGATTCACCCCTTCT 125
|||||
34 erserlleserlysarqasnseryalasnCysasnGLYTPRqYthrArg 50
|||||

126 CTTCAATCTCCAAAGGCAATAGTGTCAATTCATTCAGATGGCTGGAGAACACGA 175
 51 CysSerValAlaIalysAspTyrThrValProSerSerAlaValAspGlyG1 67
 176 TGCCTCCGTTGGCCAAAGATTACAGATTCCTCTCCAGCGGTCCGACGGCGG 225
 67 yProAlaIalagIuleuAspCysValIleValAlaIalagIylIleSerGlyL 84
 226 ACCGCGCCGGAGCTGACACTGTGTATATAGTTGGAGCCAGGAATATGTGGCC 275
 84 euCysIleAlaIalValMetSerAlaAsnTyrProAsnLeuMetValThr 100
 276 TCTCGCATTTGCCAGGTGATGTCCGCTAATATACCCCAATTTGATGTACACC 325
 101 GluAlaIarGAspAraGAlaGlyIAsnIleThrThrValGluAraGAspG1 117
 326 GAGGCGAGAGATCGTCCGGTGGCAACATAACGACTGTGGMAAGAGACGG 375
 117 YTrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134
 376 CTAATTTGGGAAAGAGTCCCAACAGTTCCAGCCGTCGATCTATATGT 425
 134 eurThrMetAlaValAspCysGlyLeuIleuIleuIleuValIleuGlyAsp 150
 426 TGACATATGGCAGTATGTGTGGATTTGAAAGATGATTTGGTGGGAGAT 475
 151 ProAsnAlaProAraGpPheValLeuTrpIleGlyIleuIleuAraGpValPr 167
 476 CCTAATATGGCCCTTCTCTGTTTGTGGAAAGGTAATTAAGCCCGTCC 525
 167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
 526 CTCAAAACCTCAGTATCTCCCTTTTGTGATTTGATGAGCATTCCTCCGCA 575
 184 yLsLeuAraGAlaGlyPheGlyProIleGlyLeuAraGpProSerProGly 200
 576 AGTTGAGAGCTGCTTTGGTCCATTTGGCCCTCCCTCAACCTCCAGGT 625
 201 HisGluGluSerValGluGlnPheValAraGAsnIleuGlyIleuGlyIleu 217
 626 CATAGGCAATCAGTTGAGCACTTGTGCTGATGATCTGGTGGCGAAGT 675
 217 lPheGluAraGpLeuIleGluProPheCysSerGlyValIleuIleuValGlyAsp 234
 676 CTTTGAACGGCTTGAADAGACCATTTTGTGCTTATGCTGATGATC 725
 234 roSerLysLeuSerMetLysAlaIalAraPheGlyIleuIleuIleuGlu 250
 726 CCTCAAAAACGAGTATGAAAGCAGCATTTGGGAAAGTTTGGAAAGTTGGAA 775
 251 GluThrGlyGlySerIleIleGlyIleuIleuIleuIleuIleuIleuIleu 267
 776 GAAACCTGGTGTACCATTTATTTGGAGAACCTTTTAAACCAATTAAGGGAG 825
 267 gSerSerThrProLysAlaProAraGAspProAraGpLeuProLysProLysG 284
 826 ATCCAGATACACCTTAAGCGCCCGCATTCGCTTACTTAACCAAAAG 875
 284 LyGIlnThrValGlySerPheAraGpIleuIleuAraGpMetLeuProAspAla 300
 876 GACGAGACGTTGGATCATTCAGAAAGGCTCCACAATATGTCGCCGAGGCA 925
 301 lIleSerAlaAraGpLeuGlySerLysIleuIleuIleuIleuIleuIleuIleu 317
 926 ATCAGTGGAAAGATTGGAGAACAAATTAACCAATATCAGGAAAGCTTTCTAG 975
 317 rIleThrLysSerGluLysGlyIleuIleuIleuIleuIleuIleuIleuIleu 334
 976 CATTAACATTAAGTACAGAAAAGAGGATATCACTTGCATACGAGACACACAG 1025
 334 lGlyValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 350

1026 AAGAGTAGTTCCTTCAAAAGTCGAAGCATTTGTCAATGACTGTCGGCATCC 1075
 351 TyrValAlaSerAsnIleLeuAraGpLeuSerValAlaIalAlaAspAl 367
 1076 TATGTAGCAAGCAACATATATAGCTCCTCTTGGGTGGCCGACGATGC 1125
 367 alaSerAsnPheTyrTyrProProValGlyValAlaValThrIleSerTyrP 384
 1126 ACTTTAAATTTCTACTATCCCACTTGGAGCAGTCAACATTTCTATATC 1175
 384 roGlnGluAlaIalAraGAspGluAraGpLeuValAspGlyIleuLysGly 400
 1176 CTCAGAACCTATTCGTGATGAGCCGTCGTTGATGCTGAACATAAAGGGA 1225
 401 PheGlyGlnLeuHisProAraGpThrGlnGlyValGluThrLeuGlyIleu 417
 1226 TTTGGGCACTTGCATCCAGCTACACAGGAGGAGGAAACCTAGAAACGAT 1275
 417 eTyrSerSerSerLeuPheProAsnAraGAlaProLysGlyAraGValLeuL 434
 1276 ATATAGTCAATCCTCTTCCCTAACCGTCCCAAAAGGTCGGGTCTAC 1325
 434 euLeuAsnTyrIleGlyIleuIleuIleuIleuIleuIleuIleuIleu 450
 1326 TCTTGAACATAATGGAGGAGCAAAAAATCCGAAATTTGTCTAAGACG 1375
 451 GluSerGlnLeuValGluValAlaAspAraGAspLeuAraGpIleuIleu 467
 1376 GAGAGCCAACTGTGGAGTATGTGATCGTACCTCAGCAAAAAATCCTAT 1425
 467 eLysProLysAlaGlnAspProLeuValAlaGlyValAraGValTyrProG 484
 1426 AAAACCCAAAGCTCAAGATCCTTGTGTGGTGGCGAGATAGCCAC 1475
 484 lAlAlaIleProGlnPheLeuValGlyHisLeuAspThrIleuSerThrAla 500
 1476 AAGCTATCCACAGTTTTGTGGTGCATCTGATCGATACCGTAAATCTACTG 1525
 501 LysAlaAlaMetAsnAspAsnGlyLeuGlyIleuPheLeuGlyIleu 517
 1526 AAAAGCTCATATGATATATATGAGGCTTGANAGGCTGTTCTTGGGGGTRNA 1575
 517 nTyrValSerGlyValAlaIleuGlyAraGpCysValGluGlyAlaIleu 534
 1576 TTAATGTCAGGATGATGATGAGGAGGATGAGGATGATGATGAGG 1625
 534 aAlaSerGluValThrGlyPheLeuSerAraGpTyrAlaIleuIleu 548
 1626 TTGCATCCGAGGTAAACAGATTTCTGCTCCGATATGCAATACAAA 1669
 seq_name: gb_p11:AB020501
 seq_documentation_block: 1647 bp mRNA PIN 06-FEB-1999
 LOCUS AB020501 Nicotiana tabacum mRNA for plastidal protoporphyrinogen oxidase,
 DEFINITION complete cds.
 ACCESSION AB020501 GI:3929921
 VERSION AB020501.1
 KEYWORDS plastidal protoporphyrinogen oxidase.
 SOURCE Nicotiana tabacum (Strain: Samsun NY) cDNA to mRNA.
 ORGANISM Nicotiana tabacum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 1647)
 Che, F. and Watanabe, N.
 AUTHORS Direct Submission
 TITLE Submitted (26-NOV-1998) to the DDBJ/EMBL/Genbank databases.
 JOURNAL Fang-Sik Che, Nara Institute of Science and Technology, Graduate
 School of Biological Sciences; Takayamachou 8916-5, Ikoma, Nara
 630-0101, Japan (E-mail: fscbe@bs.aist.nara.ac.jp.
 Tel.:81-743-72-5453, Fax:81-743-72-5459)
 REFERENCE 2 (sites)

AUTHORS Matanabe, N., Che, F. S., Iwano, M., Takayama, S., Nakano, T., Yoshida, S.
and Isogai, A.
TITLE Molecular characterization of photomixotrophic tobacco cells
resistant to protoporphyrinogen oxidase-inhibiting herbicides
JOURNAL Plant Physiol. 118 (3), 751-758 (1998)
MEDLINE 99026276
FEATURES Location/Qualifiers
source 1..1647
organism="Nicotiana tabacum"
strain="Samsun NN"
db_xref="taxon:4097"
/note="Photomixotrophic cultured cells"
gene 1..1647
"gene="NtPPOX1"
1..1647
"gene="NtPPOX1"
CDS /gene="NtPPOX1"
/codon_start=1
/product="plastidial protoporphyrinogen oxidase"
/protein_id="BA34713.1"
/db_xref="GI:3929922"
/translation="MTTPPIAHHPNIPTHOSSSPLELNRRTSPTPESSISKRSYVNC
NGMTRCVAADYVPSAVDGGPAELDCVIVGAGISLCIAOVMSANYPNLNRYTFA
RDRAGNITVTERDGYLWEGPNSEFSDPMI/TVAVDCGLDDIVLGPNNAPRYLWK
GKLRVPSKLDIAFIDLMSIIPGKLRAGFGAIGRPPSPGHESEVDEQVRENTLGEVYR
ERLLEPPSGVYAGDPSKLSMAKAPGRKWKLEBERGGSTIGGTFRAIKRSTPRAPRD
PRLRPPKQITVGRSRKGLRMLPDAISARLGSKLSLWKSISITRSEKGGYHLTETPE
GVSLQSKSIWYWPVYASNLIRPLVAAADALSNFVYVAVAVTISYPOEATRDER
LVDELKGFQDLHPRTQGVENIGTYSSSLPLPNPAPKRVYLLNVIIGAKRPILISKT
ESQLEVYVDRDLRKLIRKKAODPLLVYVWVWPAIPFVIGLHDLTSTAKAANNDSG
LEGLEFGNNYSGVALGRCEVGEVASEVSEVTFLSRAYK"

BASE COUNT 437 a 349 c 415 g 446 t
ORIGIN

alignment_scores: Quality: 2824.00 Length: 548
Ratio: 5.172 Gaps: 0
Percent Similarity: 99.635 Percent Identity: 99.453

alignment_block: US-09-508-418-2 x AB020501 ..

Align seg 1/1 to: AB020501 from: 1 to: 1647
1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSe 17
|||||
1 ATGACAAACAACCTCCGCAATCCCAATTAATTTTCACATCCACCAGTC 50
17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheS 34
|||||
51 GTGTGTCATCCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCG 100
34 exSerIleSerIlySerIlyArgAsnSerValAsnCyAsnGlyTrpArgThrArg 50
|||||
101 CTTCAATTCGCAAGCGCAATAGTGTCAATTCGCAATTCGCAATTCGCA 150
51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAlaSpGlyG 67
|||||
151 TGCTCCGTTGCCAAGAGATTACAGATTCCTCCGTCAGCGGTCCGACG 200
67 YProAlaAlaIleLeuAspCysValIleValAlaGlyAlaGlyIleSerGlyL 84
|||||
201 ACCCGCGCGGAGCTGAGCTGTATAGTGGAGCGAGGAATTAAGTGCC 250
84 euCysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
|||||
251 TCTGCATTCGCGAGGTGTCCGCTAATTTACCCCAATTTGAMTGGTAAC 300
101 GluAlaArgAspArgAlaGlyIlyAsnIleThrThrValGluArgAspG 117
|||||
301 GAGCGAGAGATTCGTCCGTCACATTAACGACACTGGGAAAGAGACGG 350
117 yTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134

||||| 351 CTAATTTGTGGAAAGAGGTCCCAACAGTTTCCAGCCGTCGATCCTATG 400
134 euThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyasp 150
||||| 401 TGACATTTGGCAGTAGATTTGGATTGGAAGATTTGGTTGGGAAAT 450
151 ProAsnAlaProArgPheValLeuTrpIlyGlyLysLeuArgProValPr 167
||||| 451 CCTAAATGGCCCGCTTTCGTTTGTGGAAAGGTAATTAAGCCCGGTCC 500
167 oSerIlyLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
||||| 501 CTCAAAACCTCACGTGATCTTGCCCTTTTGTATTGATGAGCATTCCTCG 550
184 yLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
||||| 551 AGTTGAGAGCTGAGTGTGTGTCATTTGGTCCAGCCGCTCCACCTCCAG 600
201 HisGluGluSerValGluInPheValArgArgAsnLeuGlyGlyValu 217
||||| 601 CATGAGGAATTCAGTTGAGGAGCTTCGTCGCTGATTCCTGGTGGCAAG 650
217 lPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyasp 234
||||| 651 CTTTGAACCTTGATGAAACCAATTTGTTGTTCTGTGTTATGCTGGTGA 700
234 rSerIlyLeuSerMetIlyAsnAlaIlePheGlyLysValTrpIlySerI 250
||||| 701 CCTCAAACTGAGTATGAAAGCAATTTGGGAAAGTGGAAAGTTGGAA 750
251 GluThrGlyGlySerIleIleGlyIlyThrPheLysAlaIleLysGlu 267
||||| 751 GAAACTGGTGTAGCATTATTGAGGAAACCTTTAAAGCAATTAAGAGAG 800
267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLys 284
||||| 801 ATCCAGTACACCTAAAGCCCGCCCGGATCCCGGTTTACCTAAACCAAA 850
284 lYGIlnThrValGlySerPheArgIlySGLYLeuArgMetLeuProAsp 300
||||| 851 GACACACAGTTGGATTCATCCAGGAAAGGCTTCACAGAACTGCTCCG 900
301 lLeSerAlaArgLeuGlySerIlyLeuLysIlySerIlyLeuSerI 317
||||| 901 ATCAGTGCAGAGATTGGAAAGCAATTAATAACTATCATGAGGACTTT 950
317 rIleThrIlySerGlyLysGlyIlyTyrHisLeuThrTyrGluThrPro 334
||||| 951 CMTTACTAAGTCAGAAAAAGAGATATACACTTGACATVACGACAGCAG 1000
334 lGluIlyValIleSerLeuGlnSerArgSerIleValMetThrValPro 350
||||| 1001 AAGGATGATTTCTCTTCAAAAGTCCAAACATTTGCATGATGCTGACAT 1050
351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAsp 367
||||| 1051 TATGAGGCAAGCAACATATATAGTCCCTCTTCCGTTGGCGACAGATG 1100
367 aLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyr 384
||||| 1101 ACTTTCAAAATTTCTATCATATCCCGCCAGTTGGAGCAGTCCAAATTC 1150
384 rGlnGlnAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGly 400
||||| 1151 CTCAGAAAGCTATTCGATGAGCGCTGCTGTTGATGGTGAACCTAAAG 1200
401 PheGlyIlnLeuHisProArgTyrGlnGlyValAlaGluThrLeuGlyThr 417
||||| 1201 TTTTGGCGAGTTGCATCCACGTACACAGGGAGAGTAAACACTAAGAAC 1250
417 eTyrSerSerLeuPheProAsnArgAlaProLysGlyValArgAlaLeu 434
|||||

1251 ATATAGTTCATCACTCTTCCCTTAACCGTGGCCCAAAAAGTGGGGTGTACT 1300
 434 euleuasnryrllieglyglalalysasnproglulleuSerLysThr 450
 1301 TCTTGAACCTACATTGGAGGACAAAATTCCTGAAATTTTGTCTAAGACG 1350
 451 GluSerGlnLeuValGluValValAspArgAspLeuArgLysMetLeuI 467
 1351 GAGGCCAACCTGTGGAGAGTAGTTGATCGTACCTCGAAGAAAATTCCTTAT 1400
 467 eLysProLysAlaGlnAspProLeuValValGlyValArgValTrpProG 484
 1401 AAAAACCAGAACGCTCAAGATCCTCTGTGGTGGGTCGAGGATMGCCAC 1450
 484 InAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
 1451 AAGCTATCCCAACAGCTTTTGGTGGTGCATCGATCGAAGGCTAAGTACGCA 1500
 501 LysAlaIleMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyLys 517
 1501 AAGCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1550
 517 nTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTrpGluV 534
 1551 TTATGTGTCAAGTGTACGATTTGGGGAGGCTGTGTGAAGGTGCTTATGAA 1600
 534 AlaIleSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
 1601 TTGGATCCCGAGGATTTGTCTGCGATGATGATGATGATGATGATGATGAT 1644
 seq_name: gb_p13:NTY13465

seq_documentation block: 1892 bp mRNA PLN 08-SEP-1997
 LOCUS NTY13465
 DEFINITION Nicotiana tabacum mRNA for protoporphyrin IX oxidase, ppxI gene.
 ACCESSION Y13465
 VERSION Y13465.1 GI:2370332
 KEYWORDS oxidase; ppxI gene; protoporphyrin IX; tetrapyrrole synthesis enzyme.
 SOURCE common tobacco.
 ORGANISM Nicotiana tabacum
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; easterids I; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 1892)
 Lemontova, I., Kruse, E., Mock, H.P. and Grimm, B. Cloning and characterization of a plastidial and a mitochondrial isoform of tobacco protoporphyrinogen IX oxidase Proc. Natl. Acad. Sci. U.S.A. 94 (16), 8895-8900 (1997)
 JOURNAL MEDLINE 97385200
 REFERENCE 2 (bases 1 to 1892)
 Inna, L. Direct Submission
 Submitted (29-MAY-1997) L. Inna, Institute of Plant Genetics (IPK), Dept. of Molecular Cell Biology, Corrensstr.3, 06466 Gatersleben, FRG
 REMARK revised by author
 FEATURES
 source location/Qualifiers
 1. 1892
 /organism="Nicotiana tabacum"
 /strain="SRI"
 /db_xref="taxon:4097"
 /clone_lib="Lambda ZAP II tobacco library"
 29. 1675
 /gene="ppxI"
 29. 1675
 /gene="ppxI"
 /EC_number="1.3.3.4"
 /function="tetrapyrrole biosynthesis"
 /codon_start=1
 /product="protoporphyrinogen oxidase"
 /protein_id="CAA73865.1"
 /db_xref="GI:2370333"

/db_xref="SWISS-PROT:O24163"
 /translation="MTTTPPIAHNPPIFTHOSSSSPLAFLNRTSFTFESSISKRNSVNC
 NGRFRCSVAKQDYVPSAIVGGPAAELDVCVAGISGLCIAQVMSAVYPLMVTTEA
 RDRAGGNTTIVEDQYLMEGRPSRPPSPMIFMNVDCIKDILVIGDNNARFVLMK
 GKLRVPSKTIIDLAFDLMNIRKLRAGGALIGRSPFGHESVDEQFRRRLGGEVF
 ERLEPFGSGVYADGDPKLSMKRAFKGVKLEBTDGSLIGTFFKAIKERSYKAPRD
 PRLPKRQGVVSGFSRRKGLMLPDALISARLGSIKLKSMLKLSITKSEGGYHLTYTPE
 GYVLSQRSIVMTVPSYVASNILRPLSVAADALSNFYYPYGAVTITYPDAIBDER
 LVDPELTKFQLEHPTROGVEETLGTIVSSLEFPRARPKGVLLNLYIGCAKNEEISKT
 ESDLVPEVBDLLEKMLIKPKAODPLVGVYVMVQAIPQFLVGHLDLSTAKAAMDNDG
 LBSFLGNTVSVVALGRVDEGAYEVASEVYTGFLSRAYAK"
 BASE COUNT 516 a 383 c 467 g 526 t
 ORIGIN

alignment_scores: Length: 548
 Quality: 2821.00 Gaps: 0
 Ratio: 5.167
 Percent Similarity: 99.635 Percent Identity: 99.270

alignment_block:
 US-09-508-418-2 x NTY13465

Align seg 1/1 to: NTY13465 from: 1 to: 1892
 1 MetThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSe 17
 29 ARGACACACAGCTCCATCCATCATCCTAATATTTTCACACACACAGCTC 78
 17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheS 34
 79 GTCGTCATCCGCAATGGCCATTCCTTAACCGCTACGAGTTCATCCCTTCT 128
 34 eRSerIleSerLysArgAsnSerValAsnCysAsnGlyTyrPArgThrArg 50
 129 CTTCAATCTCCAAAGCCCAATATGTCATATGCAATGGCTGGAGAACAGCA 178
 51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAspGlyG 67
 179 TGCTCCGTTGCCAAGATTAACACAGTTCCTCCACAGCGGTCGACGGCGG 228
 67 yProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyL 84
 229 ACCCGCGCGGAGCTGGACTGTGTATTAAGTGGAGCGAAGTAAAGTGGCC 278
 84 eucYsIleLeuAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
 279 TGTGCAATTTGGCGCAGTGGATGCTCCGCTAATTAACCCCAATTTGATGGTAACC 328
 101 GluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgAspGl 117
 329 GAGCGCAGAGATGCTCCGCTGGCAATTAACGACTGTGGAAAGAGACGG 378
 117 yTyrLeuTTPpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134
 379 GYATTTGTGGGAGGAGGTCCTCAACAGTTTCCAGCGCTCCGATCTATGAT 428
 134 euthrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAsp 150
 429 TGTACTATGGCAGTACGATTTGGATTAAGAGATGATTTGGTGGAGAGAT 478
 151 proAsnAlaProArgPheValLeuTyrPlyGlyLysLeuArgProValPr 167
 479 CCTAATGCGCGCCCGTTGCTTTGTGGAAGGTTAAATTAAGCCCGCTCC 528
 167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyYL 184
 529 CTCMAAACCTCACGATCTGCTTTTATTTGATGACCAATTTCCCTGGGA 578
 184 ySLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
 579 AGTTGAGAGGCTGTTTGGTGGCATTGGCCCTCCGCCCTTCCACTCCAGGT 628

201 HlsgLInGluSerValGluInpHeValAARgArqsnLeuGIyGIyLuVA 217
 629 CARGAAGAAACAGTGTGAGCAGGTTCGTCCGCGTAAATCTGTGGCCAGACT 678
 217 lPheGluAqGLeuIlleGluProPheCySerGIyAlTYrValGIyAspP 234
 679 CTTTGAACCGTTGATGAAACCAATTTTGTCTGGTGTATGCTGGGGATC 728
 234 roserLySerMetLysAlaAlaPheGIyLysValTrpLysLeuGI 250
 729 CTTCAAAACTGAGTATGAAAGCAGCAATTTGGAAAGTTGGAAAGTTGGAA 778
 251 GluTrhGIySerIlleGIyGIyTrhPheLysAlaIlleLysGIuAr 267
 779 GAACAGGTGGTGTAGCAFTTATGGAGAACCTTTAAAGCAATAAAGAGAG 828
 267 gSerSerTrhProLysAlaProArqAspProArqLeuProLysProLysG 284
 829 ATCCAGTACACTTAAAGCCGCCCGCATCCGCTTAAACCTAAACCAAAAG 878
 284 lYGIInPheValGIySerPheArqLysGIyLeuArqMetLeuProAspAla 300
 879 GACAGACAGTGTGATCATTCAGAGAGGCTCTCAGAAATGCTGCCGGATGCA 928
 301 lLSeValArqLeuGIySerLysLeuLysLeuSerTrpLysLeuSerSe 317
 929 ATCAGTGAAGATTTGGGAAGCAAAATTAACAATCAATGGAAGCTTTCTAG 978
 317 rIleTrhLysSerGIyLysGIyLysTrhLysLeuTrhTYrGIuTrhProG 334
 979 CATTACTAACTCAGAAAAGAGAGATATCACTTACATCAGAACACACCCAG 1028
 334 lInGIyValLysTrhLeuGIySerArqSerIlleValMetTrhValProSer 350
 1029 AAGCAGTAGTCTCTTCAAGAGTCGAAGCAATTTCTCATGACTGGCCATCC 1078
 351 TYrValAlaSerAsnIleLeuArqProLeuSerValAlaAlaAlaAspAl 367
 1079 TATGTACCAAGCAAAATATTAAGTCCCTTGGTGGCCACAGCAGATGC 1128
 367 aLeuSerAsnPheTYrTYrProProValGIyAlaValTrhIleSerTYrP 384
 1129 ACTTCAAAATTTCTACTATCCCCAGTGGAGCAGCACCAATTAACAATATC 1178
 384 roGIInGIuAlaIleArqAspGIuArqLeuValAspGIyGIuLeuLysGIy 400
 1179 CTCAGAAAGCTATTTCTGATGAGCGCTGTGTATGATGGTAACTAAAGGGA 1228
 401 PheGIyLInLeuHisProArqTrhTrhGIyValGIyTrhLeuGIyTrhIl 417
 1229 TTTGGGCAAGTTGGATCCACGCTACACAGGGAGTGGAAACACTAGGAACGAT 1278
 417 eTYrSerSerSerLeuPheProAsnArqAlaProLysGIyArValIleuL 434
 1279 ATAAAGTTCATCTCTCCCTTAAACCGTCCCAAAAAGGCGGGTCTTAC 1328
 434 euLeuAsnTYrIleGIyAlaLysAsnProGIuIleLeuSerLysTrh 450
 1329 TCTTGAACCTCATTTGGAGGAGCAAAAATAAATCTGAAATTTTGTCTAAAGACG 1378
 451 GluSerGIuLeuValGIyValAspArqAspLeuArqLysMetLeuIl 467
 1379 GAGAGCCAACTTGTGGAAAGTATGATCGTACCTCGAATAAATGCTTAT 1428
 467 eLYsProLysAlaGIuAspProLeuValAlaGIyValArqValTrpProG 484
 1429 AAAACCCAAAGCTCAAGATCCTCTTGTGGTGGTGGAGTAAAGGCCAC 1478
 484 lAlaIleProGIuInPheLeuValGIyHisLeuAspTrhLeuSerTrhAla 500
 1479 AAGCTATCCACAGATTTTGGTGGTGGTGGATCGATACGCTAAAGTACTGCA 1528
 501 LysAlaIleMetAsnAspAsnGIyLeuGIyLeuPheLeuGIyGIyAs 517

1529 AAGCTGTATGAATGATTAATGAGCTTGAAGGCTGTCTTGGGGTAA 1578
 517 nTYrValSerGIyValAlaLeuGIyArqCyAsValGIuGIyAlaTrpGIuV 534
 1579 TTATGTGTACAGGTGTAGCATTTGGGAGGAGTGTGTGAAGGCTTATGAG 1628
 534 aAlaSerGIuValTrhTrhGIyPheLeuSerArqTYrAlaTYrLys 548
 1629 TTGCATCCGAGGTAAACAGATTTGCTCGGTATGCAATACAAA 1672
 seq_name: gb_P14:STA225107

seq_documentation_block:
 STA225107 1976 bp mRNA PUN 09-FEB-1999
 DEFINITION Solanum tuberosum (cultivar Bintje) chloroplastic
 protoporphyrinogen IX oxidase.
 ACCESSION AJ225107
 VERSION AJ225107.1 GI:3093409
 KEYWORDS protoporphyrinogen oxidase.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
 I; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 1976)
 Johnston, D. J.

REFERENCE Submitted (23-APR-1998) Johnston D. J., Biologie moleculaire, RAC,
 Changins, 1260, NYON, SWITZERLAND
 2 (bases 1 to 1976)
 Johnston, D. J., Droz, E., Rochaix, J. D. and Malno, P.
 Cloning and Characterization of Potato cDNAs Involved in
 Tetrapyrrole Biosynthesis: ferrochelatase (Accession No. AJ005802),
 chloroplastic protoporphyrinogen IX oxidase (Accession No.
 AJ225107), and mitochondrial protoporphyrinogen IX oxidase
 (Accession No. AJ225108)
 Plant Physiol. 118, 330-330 (1998)
 Location/Qualifiers
 1..1976
 /organism="Solanum tuberosum"
 /cultivar="Bintje"
 /db_xref="taxon:4113"
 1..1674
 /EC_number="1.3.3.4"
 /function="tetrapyrrole biosynthesis"
 /codon_start=1
 /product="protoporphyrinogen oxidase"
 /protein_id="CAA12400.1"
 /db_xref="GI:3093410"
 /db_xref="SPRREMBL:064384"
 /translation="MTTAVANHSIFTHRSPPSPSSSSSPFFLNRNFTIPYS
 TSKRNSVNCWGRKRCVAKDYVTPSPSEVDNGNFPPLDCVYVAGISGLCIKXVIGAN
 YPILNMTYEAARDRAGGNTTYERGGYIAMBESPNFQSDMLTKAANDCGIKDDLYIGDP
 DAREFVIAWKDKLVPKGLTDLPFDLMTSPKLRAGFALIGLRBSPGIGESBEDQF
 RRMNGAWEFERLLEPFCGVYAGDPKRLIKRAFAFGYWKLEDTGSGSIIIGSTRKAIKR
 SNMPKPPDRDLRPLPKGOYVGFSEFKGLRMLPDAICRLSGSKYLSWKLSSIRKSEKGG
 YLLHYEPBEVSVLSRSRYVMVSVYVAISMLRPLSVAADALSFRYYPVLAAVTISY
 PORAIRBERLYDDELKQHPRSQGVVETLIGTIVSSSLFPNBARAPGKRVLLNLYIGGA
 TNRTEIYSKTSQGLFPEAVDRDLRKMILKRAQDEPFYVGVYVWVWQALTPQPLVGHLDLIGT
 AKTALSDNGJDLFLEGGNYVSGVALRCKVBAVEIASVETGFLSDYATK"

JOURNAL FEATURES source
 CDS
 /ORIGIN

3' UTR
 BASE COUNT 518 a 394 c 485 g 579 t
 ORIGIN

alignment_scores:
 Quality: 2565.50 Length: 557
 Ratio: 4.850 Gaps: 2
 Percent Similarity: 94.973 Percent Identity: 88.869

alignment_block:
 US-09-508-418-2 x STA225107

Align seg 1/1 to: STA225107 from: 1 to: 1976

1 MetPrHrThrProIleAlaHisProAsnIlePheThrHisGln.. 16
 1 ATGACAACAGCGCCGCGTCCACATCCATCCAGCTACCTCACCGGTC 50
 17SerserSerserPro...LeuAlaPheL 25
 51 GCCCGTCCGCGTCCGCGTCCCTCCCTCCGCGCATTTTATTTT 100
 25 euAsnArgThrSerPheIleProPheSerIleSerLysArgAsnSer 41
 101 TAAACCGTACCAATTTTCATTCCTTACCTTCCACCTCCAAAGCGCAATVGT 150
 42 ValAsnCysAsnGlyTrpArgThrArgCysSerValAlaLysAspTyrTh 58
 151 GTCATATTGCATGGCTGGAGAACACGATGTTCCGTTGCCAAGAGATTAAC 200
 58 rValProSerSerAlaValAspGlyGlyProAlaAlaGluLeuAspCysV 75
 201 AGTTCCTCCCTCGGAAGTCCGACGGTAAATCAAGTCCCGGAGCTGGATTGTG 250
 75 allIleValGlyAlaGlyIleSerGlyLeuCysIleAlaGlnValMetSer 91
 251 TGGTAGTTGGAGCAGGAATTAGTGAGCTGCAATGGCTAAAGTAAATTCG 300
 92 AlaAsnTyrProAsnLeuMetValThrGluAlaArgAspArgAlaGlyL 108
 301 GCTAATATTCCTCCAAATTTGATGTGACGAGGAGCGGAGATCCTGCCGGTGG 350
 108 yAsnIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProA 125
 351 AAACATAACGACGGTGGAAAGATGATCTTATGCGAAAGAAAGCTCTTA 400
 125 snSerPheGlnProSerAspProMetLeuThrMetAlaValAspCysGly 141
 401 ACAAGTTTCCAGCCTTCCGATCTTAATGTTGACAAATGGCTGTAGATTTGGA 450
 142 LeuLysAspAspLeuValIleuGlyAspProAsnAlaProArgPheValI 158
 451 TTGAAGAGTATTTGGTGGGAGATCCTGATCGCCCTGGCTTTGGTCTT 500
 158 uTrpLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProP 175
 501 GTGGAAGGATTAACCTAAGCCCTGTTCCCGCAACCTCACATCTTCCCT 550
 175 hePheAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyPro 191
 551 TCTTTGATTTGATGATGATATCCCTGGCAAGCTCAGAGCTGGTTTGGTGGC 600
 192 IleGlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPh 208
 601 APTGGCCCTGCGCCCTTCCACCTCCAGGTTATGAGAAATCAGTTGAGCAGTT 650
 208 eValArgArgAsnLeuGlyGlyGluValIlePheGluValLeuIleGluProP 225
 651 CGTCCGTCGTAAATCTTGTGTCAGAAAGCTTTTGAACCGTTTGTATGAACCAT 700
 225 heCysSerGlyValTyrValGlyLysProSerLysLeuSerMetLysAla 241
 701 TTTCTTCTGGTGTTCACCGCGGAGCCCTCAAAATTTGATTAAGAAAGCA 750
 242 AlaPheGlyLysValTyrLysLeuGluGluThrGlyGlySerIleIleGle 258
 751 GCATTTTGGAAAGTGTGAGAGCTAAGAAACAATGCTGATTTATTTGG 800
 258 yGlyThrPheLysAlaIleLysGlyLysArgSerSerThrProLysAlaProA 275
 801 GGGAAACCTTTTAAAGCAATTAAGAGAGATCCAGTAAACCTCCAAACCCCTCC 850
 275 rGAspProArgLeuProLysProLysGlyGlnThrValGlySerPheArg 291

851 GTGATCCCGCGTTTACCAACACCAAAAAGGACAAACCTGTGGATCATTTAGG 900
 292 LysGlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLys 308
 901 AAGGGTCTGGAATGCTGCCGAGTCCAAATTTGTAAGAAAGCTGGAAAGCA 950
 308 sLeuLysLeuSerTrpLysLeuSerSerIlePheLysSerGlyLysGly 325
 951 AGTAAACATTCATTCGAAAGCTTTCTAGCATTTACAAAGTACGAAAAAGAG 1000
 325 LyrTrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSer 341
 1001 GATATCTCTGACATPACGAGACACACGAAAGAGTACTTCTCGCGAAGT 1050
 342 ArgSerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArg 358
 1051 CGAAGCATTTGTCATGACCTGTCCATCTTAATGTAAGCAACATATPACG 1100
 358 pProLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProP 375
 1101 CCTCTTTCCGGTCCGTCGACAGCATCACTTCAAGTTTCTACTAATCCCC 1150
 375 roValGlyValAlaThrIleSerTyrProGlnGluAlaIleArgAspGlu 391
 1151 CAGTACACACAGTACAAATTTTATATCCCAAGAGCCTATTCGTGATGAG 1200
 392 ArgLeuValAspGlyLysGlyPheGlyGlnLeuHisProArgThr 408
 1201 CGCTGTGTTGATGATGTAACSTAAAGGATTTGGGCAAGTTCATCCACGTTT 1250
 408 rGlnGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProA 425
 1251 ACAGGAGAGTGAACACACTAGAAACAATATATATGTTTCAATCCTTTCCCTA 1300
 425 snArgAlaProLysGlyArgValIleLeuLeuAsnTyrIleGlyGlyAla 441
 1301 ACCGTCTCCAAATGGCCGCGGTTGCTACTTGAACCTCACTAATGGAGGACCA 1350
 442 LysAsnProGluIleLeuSerLysThrGlnSerGlnLeuValGluValVa 458
 1351 ACAAAATACGAAATTTGCTGTAAGACGGAGACCAACTTTGTGGAAAGCAGT 1400
 458 LAspArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProL 475
 1401 TGAACCTGACCTCAGAAAATGCTTATTAACCCAAACACAAAGATCCCT 1450
 475 euValValGlyValAlaArgValTrpProGlnAlaIleProGlnPheLeuVal 491
 1451 TTGTTTAGGGGTGTGCGAGTATGCGCAAGCTATCCCAAGTTTTGGTTC 1500
 492 GlyHisLeuAspThrLeuSerThrAlaLysAlaIleMetAsnAspAsnG 508
 1501 GGACATCTGTATACACTAGGTACTGCAAAAACCTGCTCAAGAGATTAATGG 1550
 508 yLeuGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuG 525
 1551 GCTTGGACGGGCTAATTCCTTGGGGGTAAATATGCTGCTGTGATGAGCATGG 1600
 525 LArgCysValGluGlyAlaIleArgIleValAlaSerGluValThrGlyPhe 541
 1601 GAAAGTGTGTTAAAGTCTTATGAAATATGACATCTGAGAGTAACTGGATTT 1650
 542 LeuSerArgTyrAlaTyrLys 548
 1651 CTGTCTCAGTATGACATACAAA 1671

seq_name: gb_pat2:AX084746
 seq_documentation_block: 1826 bp DNA PAT 09-MAR-2001
 LOCUS AX084746 DEFINITION Sequence 15 from Patent W00112825.
 ACCESSION AX084746 VERSION AX084746.1 GI:13274999

1380 TCGTATTTGAGAAAAATGCTTATTAATCCATAATGCAAAAGATCCCTCTTG 1429
 476 aVAlaGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
 1430 TTTTGGGCTCTAGAGATATAGCCAAAGCATTCCACAGTTCTTGSTTGGT 1479
 493 HIsLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
 1480 CATTGGATCTCCTTGATAGTGCACAAAATGCGCTCTCAGGGATTTGGGTT 1529
 509 uGlnGlyLeuPheLeuGlyGlyAsnTrpValSerGlyValAlaLeuGlyA 526
 1530 TCATGGACGTCTTCTTGGGGCACTATGATCGTGTGTGCATTAAGGAC 1579
 526 TGCysValaGlnGlyAlaTrpGluValAlaSerGluValThrGlyPheLeu 542
 1580 GGTGTGTGGAAAGTGTCTTACCGAGGTGGACCTGAAGATTAAGAGAAATTC 1629
 543 SerArgTyrAlaTyrLys 548
 1630 TCACAATATMGCAATACAAA 1647

seq_name: gb_Dat1:A78957
 seq_documentation_block: 1704 bp DNA PAT 20-OCT-1999
 LOCUS A78957 Sequence 11 from Patent WO9829554.
 DEFINITION A78957
 ACCESSION A78957
 VERSION A78957.1 GI:6092085
 KEYWORDS
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicotls; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 (bases 1 to 1704)
 AUTHORS Ishige, F., and Sato, R.
 TITLE Ishige, F., and Sato, R.
 JOURNAL METHODS OF CONFERRING PRO-INHIBITING HERBICIDE RESISTANCE TO PLANTS
 BY GENE MANIPULATION
 Patent: WO 9829554-A 11.09-JUL-1998;
 IShIGE FUMIHARU (JP); SAHO RYO (JP)
 FEATURES
 source Location/Qualifiers
 1..1704 /organism="Arabidopsis thaliana"
 /strain="ECOTYPE COLUMBIA"
 /db_xref="taxon:3702"
 16..1629 /codon_start=1
 /product="PROTOPORPHIRINOGEN OXIDASE"
 /protein_id="CA59132.1"
 /db_xref="GI:6092086"

CDS
 /translation="MELSLRPTTOSLSPFSKPNLRNLNVYKRLRRCSSVAGGPTVGS
 /trnaslation="MELSLRPTTOSLSPFSKPNLRNLNVYKRLRRCSSVAGGPTVGS
 SKTEGGGTTITTCVYIVGGTSGCINQALATFKHPDAPNLIVAEADRVNGNITTR
 ENRGFLMEGSPNSFQSPDMLTMVYDSSGLKDDLYAGDPAPRFVYVANGKLRVPSKLT
 DLFEDLMSIGKIRAGFALGIRSPPEGRSVEFERNLNGIDVFEKLELPFESKLV
 YADPDKLSMKAAFKVWRLEQNGSIIIGTQKAIQERNKAPARBDPLPPOQTV
 GSRKGLRMIPKAIARLGSKYKLSWIKLSGIVKLSGYNLLETDPGLVSVOSKSV
 MYVPSHVASGLRPLSESAANALSKLYPPVAASISYKELIRPECLIDGELKGGEG
 LHRPQGVETLGTIYSSSLFVFNRAAPGRILLNLYGGSTNNGIISKSSGELVYEAARD
 LKMLIKRPSPTPIKIGVRRVWPAQIPQLVGHFDLIDPRKSSLISSGTEGLFLGSNVY
 AGVALGRCEVGEYETAIENVMNSRAIYK"
 BASE COUNT 453 a 348 c 431 g 472 t
 ORIGIN
 Alignment_scores:
 Quality: 2077.00 Length: 539
 Ratio: 4.230 Gaps: 4
 Percent Similarity: 91.095 Percent Identity: 75.139
 Alignment_block:
 US-09-508-418-2 x A78957

Align seq 1/1 to: A78957 From: 1 to: 1704

19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
 13 TCATGGAGATTATCTCTCCGCGACGACTCAATCCGCTTCTCCGTC 62
 35 ILeSerLysArgAsnSerValAsnCysAsnGlyTrp.....ArgT 49
 63 GFTTTCGAACCCCAAT...CTCCGATTAAATGTTTMAAGCCCTTAGAC 109
 49 hArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
 110 TCCGTGTTCAGTGGCCGGTGGACCAACCTCCGACTCTCAAAAACGGAA 159
 66 GlyGlyProAlaIleGluLeu.....AspCysValIleValGlyAlaG 80
 160 GCGCGAGAGGACACCACCAATCAGCAGCGATTTGTATGTTGCGCGAGG 209
 80 yLeSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTrp..... 94
 210 TATTAGTGTCTTGGCATCCGCTCAGCGCGCTTGCTACTAAGCATCCGATG 259
 95ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyIAsn 109
 260 CTGCTCCGAAATTAATTTGATCCAGCGCTVAAGATGCTGTGGAGGCAAC 309
 110 IleThrThrValGluArgAspGlyTyrLeuTrpGlnGlyProAsnSe 126
 310 ATATATCACTCGTGAAGAAATGTTTCTCTGGGAAAGAGTCCCAATAG 359
 126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 143
 360 TTTTCAACCGTGTGATCCATATGCTACTATMGTRAGATAGATGTTTGA 409
 143 yAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
 410 AGGATGATTTGGTGTGGAGATCTACTGCGCCCAAGGTTTGGTGTGG 459
 160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176
 460 AATGGGAATATGAGCGCGGTTCCATGAAAGTAAAGACTTCCCGCTT 509
 176 eaSPLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
 510 TGAATTTGATGATAGTATTGGTGGAAAGATTAGAGCTGTTTGGGCACTTG 559
 193 LysLeuArgProSerProProGlyHisGlnSerValGlnIlePheVal 209
 560 GCATTTGACCCGTCACCTCCAGGTCGGAAGAAATCTGTGGAGGATTTGTA 609
 210 ArgArgAsnLeuGlyGlyGluValPheGluValArgLeuIleGluProPhe 226
 610 CCGCGTAACCTCGGTATGATGAGTTTGGAGCCGCTGATTTGAAACCGTTTGG 659
 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaIle 243
 660 TTCAGGTTTATGCTGTGGATCCCAAACTGAGCATGAAAGCAAGCGCT 709
 243 hGlyLysValTrpLysLeuGlnGluThrGlyLysIleIleGly 259
 710 TTGGGAAGGTTTGGAAACRAGAGCAAAAGTGGAAAGCTAATTAAGTGTGT 759
 260 ThrPheLysAlaIleLysGlnArgSerSerThrProLysAlaProArgAs 276
 760 ACTTTTAAAGCAATTCAGAGAGGAAAAAACGCTCCCAAGGCAAGAACGAGA 809
 276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
 810 CCCGCCCTGCCAAAACACAGGCGCCAAACAGTGTCTTCCAGAGAGG 859
 293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
 860 GACTTTCGAATGTTGCCAGAAATATCTGCAAAAGATTAAGTGGCAAAAGTT 909

310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyIle 326
 910 AAGTGTCTTGGAAAGCTCTCAGATACCACTAAGCTGAGAGCCGAGATA 959
 326 rHSLstLrThrTrpGluThrProGluGlyValAlaSerLeuGlnSerArG 343
 960 CAACCTTAACTAAGACTCCAGAGTGGTTTGGTTTCCGTCGACAGCAAAA 1009
 343 erIleValMetThrValProSerTrpValAlaSerAsnIleLeuArGPro 359
 1010 GGTGTAAATAGAGGCGGTCATCTCAGATGTCAGATGCTGCTTGGCCCT 1059
 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTrpTrpProProVa 376
 1060 CTTCGAAACCTGCTGCAATGACACTCTCAAAACCTATTATTACCCAGAGT 1109
 376 IGLValAlaValTrHLeSerTrpProGlnGlnAlaIleArgAspGluArG 393
 1110 TGCAGCAGATACACTCGACCCGAAAGCAAGCCGAAACCGAAGCAATGTT 1159
 393 euValAspGluLeuLysGlyPheGlyGlnLeuHISProArGTrpGln 409
 1160 TGAATAGATGGTGAACCTAAAGGGGTTTGGGCAATTTGATCCAGCCAGCA 1209
 410 GLYValGluThrLeuGlyTrHLeTrpSerSerSerLeuPheProAsnAr 426
 1210 GGAGTGGAAACATTAGCAACTAATCTCAGCTCTCACTCTCCCAAAATCG 1259
 426 gAlAProLysGlyArgValIleLeuLeuAsnTrpIleGlyIleAlaLysA 443
 1260 CGACCCGCGCCGAAAGATTGCTGTGGTAACCTAATTTGGCGGGCTCAACA 1309
 443 snProGluIleLeuSerLysTrHLeuSerGlnLeuValGluValAlaSp 459
 1310 ACACCCGAAATTCTGCCAAGTCTGAAGGTGAGTTAAGTGGAGAGCAATGAC 1359
 460 ArgAspLeuArGlySMetLeuIleLysProLysAlaGlnAspProLeuVa 476
 1360 AGAGATTGAGAAATGCTAAATTAAAGCCTAAATTCGACCCAGATCCACTTA 1409
 476 IValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyAr 493
 1410 ATTAGAGATTAGGCTATGGCCCTCAAGCCATTCTCAGTTTCAAGTTGGTTC 1459
 493 isLeuAspTrHLeuSerTrHAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
 1460 ACTTGTATACCTTGACACGGCGTAAATCAACTAAAGCTTCCGGGCTCAC 1509
 510 GLUgLYleuPheleuGlyLysAsnTrpValSerGlyValAlaLeuGlyAr 526
 1510 GAAGGGCTATTTTGGGTGCAATTCGCTGGTGTAGCCCTTAAAGCCGG 1559
 526 gCySValGluGlyAlaTrpGluValAlaSerGluValTrHTrpLysPheLeu 543
 1560 GTGCTGTAGAAAGGCGCATATGAACCCGCAATTAAGGCTCAACAATTTCA 1609
 543 eTrArGTrpAlaTrpLys 548
 1610 CACGGTAAACGCTTAACAAG 1626

seq_name: gb_pat2:AX084732
 seq_documentation_block:
 LOCUS AX084732 1719 bp DNA AX084732
 DEFINITION Sequence 1 from Patent W00112825.
 ACCESSION AX084732
 VERSION AX084732.1 GI:13274985
 KEYWORDS
 SOURCE thale cress:
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

REFERENCE 1 (bases 1 to 1719)
 AUTHORS Johnson, M.A., Volzath, S.L., Helfetz, P.B. and Law, M.D.
 TITLE Herbicide-tolerant protoporphyrinogen oxidase
 JOURNAL Patent: WO 0112825-A 1 22-FEB-2001;
 PARTICIPATIONS Ag (CH)
 FEATURES source location/Qualifiers
 ORIGIN 1..1719

CDs /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 31..1644
 /note="Arabidopsis protox-1"
 /codon_start=1
 /protein_id="CAC34008.1"
 /db_xref="GI:13274986"
 /translation="MELSLRSLRPTQSLRPSFSPKPNLRNVYRKLRLKRSVAGGPRVGS
 SKTEGGGTLITTDCTIVYGGISGLTIAQLAQLATKHPDAAPNLIYBAKRKRGVNIITR
 EENGFIMEBSFQPSDPMILWVDGSLKDLVLDPAAPVILWNGRLRVPKSLT
 DLPFDLMSITGKRIRAGFGALGIRPSPPEKRSVEFVRVNRNGDEVEFRLIRPFCGTV
 YADDPKRLSMKAAFGVMKLEONGCSIIGTFRKAIQERAKRABERDPLPKPOGQTV
 GSPRKGLRMLPEALISARLQSKYKLSKLGITRLEGGYNLYEPDPKLVSYOSRKY
 MTPSHVASGLRSLRSESAANLSKTYVPPVAAYSIYKPKAIPRECLTIDGELKRFQ
 LHRRTQVETLSTIYSSIFPPRPAEPGRILLNLTIGSTNWTGILSKSESELYEAVDRD
 LKRMILKPNSTDLEKLVKRWMPQALIFQFIVGHFDLIDTAKSSLFTSSGVEGLPLGGNVY
 AGVALGRVGEAVETVALEVNEMSRVAYR"

alignment_scores:
 Quality: 2077.00 Length: 539
 Ratio: 4.230 Gaps: 4
 Percent Similarity: 91.095 Percent Identity: 75.139
 alignment_block:
 us-09-508-418-2 x AX084732 ..
 Align seg 1/1 from: 1 to: 1719

19 serSerProLeuAlaPheLeuAsnArGThrSerPheIleLeuProPheSerSe 35
 28 TCCATGAGATTATATCTTCTTCCGCCGACGACTCAATCCGCTTCTTCGGTC 77
 35 rIleSerLysArGAsnSerValAsnCySAsnGlyTrp.....ArGT 49
 78 GTTTCGAAAGCCCAAT...CTCCGATTAAATGCTTAAATCAAGCCCTTCAAG 124
 49 hrArgCySerValAlaLysAspTrpValProSerSerAlaValAsp 65
 125 TCCGTTGTTCAAGTGGCCGCGTGGACCACCCTCGGATCTCAAAATCGAA 174
 66 GLYgLYpRoAlaAlaGluLeu.....AspCYValIleValGlyAlaGl 80
 175 GCGGAGAGAGGCACCCACCATCACAGAGGATTTGTGATTTGCGGGCGGAGG 224
 80 yLleSerGlyLeuCySIlleaGlnValMetSerAlaAsnTrp..... 94
 225 TATTAGTGTCTTTGGCATTCCCTCAGCGGCTTGCATAAAGGATCCGTGATG 274
 95ProAsnLeuMetValThrGluAlaArGAspArGAlaGlyLysn 109
 275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAAGGATCGTGTGGGCGCAAC 324
 110 IlleThrTrValAlaGluArGAspGlyTrpLeuTrpGluGluGlyProAsn 126
 325 ATTATACACTCTGTGAAGAAATGTTTCTCTGCGGAAGAAAGGTCCCAATAG 374
 126 rPheGlnProSerAspProMetLeuTrHMetAlaValAspCySgLYleu 143
 375 TTTTCAACCCCTCATGATCTGACTATGAGGTGTGATVAGATTGTA 424
 143 ysAApAspLeuValLeuGlyAspProAsnAlaProAlaGpPheValLeuTrp 159

425 AGCATGATTTGGTGTGGAGATTCCTACGCGCCAAAGGTTGGTTGTGG 474
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176
475 AATGGGAAATTTGAGCCGGTCCATCGAAGCTTAACAGACTTACCGTTCTT 524
176 eaSpLeuMetSerLeuProGlyLysLeuArgAlaGlyPheGlyProIleG 193
525 TGATTTGATAGTATTTGGTGGAGATTAGAGCTGGTTTGGTGGCACTTG 574
193 LysLeuArgProSerProProGlyHisGlyLeuSerValGluGlnPheVal 209
575 GCATTTGCAGCCGTCACCTCCAGGTCGTGAGAAATCTGTGGAGGATTTGTA 624
210 ArgArgAsnLeuGlyGlyValPheGluArgLeuIleGluProPheGly 226
625 CGGCGTAACTCCGGTGTGAGGTTTGGAGCCGCTGATTAACCCGTTTTCG 674
226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 243
675 TTCAGGGTGTATCTGTGTCTCTCAAAAATGAGCATGAAAGACAGCGGT 724
243 heGlyLysValTyrPheLysLeuGluGlnThrGlySerIleIleGlyGly 259
725 TTGGGAAGGTTTGGAAACTGAGGCAAAAATGTTGGAAACATTAATAGTGT 774
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg 276
775 ACTTTTAAAGCAATTCAGGAGAGGAAAAACGCTCCCAAGCAGAACGAGA 824
276 ProArgLeuProLysProLysGlyGlnThrValGlySerPheArgPhe 293
825 CCCCGCCGTCGCAAAAACACACAGGCCCAACAGTTGTTCTTTCAGAGAG 874
293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
875 GACTTCCGAAATGTTCCAGAAAGCAATATCTCCAGATTTAGGAGCAAAAGTT 924
310 LysLeuSerTyrPheLysLeuSerIleThrLysSerLysGlyGlyGly 326
925 AAGTGTGCTTGGAAAGCTCTCAGGTATCACTTAAGCTGAGACCGGAGGATTA 974
326 HisLeuThrTyrGlnThrProGluGlyValIleSerLeuGlnSerArg 343
975 CAACTTAAACATATGAGACTCCAGATGGTTTATAGTTCCTCCGAGAGCAAAA 1024
343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
1025 GTGTGTGAAATGACCGTCCATCTCATGTGTGCAAGTGTCTTCCGCGCCCT 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVal 376
1075 CTTTCTGCAATCTGCTGCAAAATGCACTCTCAAAAACATATATTTTACCCACCA 1124
376 IGLYALAVAlThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
1125 TGCAGCAGATATCTATCTATCTGTCAGCCGAAAGAACCAATCCAGAAATGTT 1174
393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgTyrGln 409
1175 TGAATGAAAGTGAAGTAAAGGTTTGGGCAATTTGATCCACCGCACGCAAA 1224
410 GLYVALGluThrIleuGlyThrIleTyrSerSerSerLeuPheProAsn 426
1225 GGAAGTTGAAATATTAAGAACTATTTACAGCTCTCTTCCAAATTCG 1274
426 gaLAProLysGlyArgValLeuLeuLeuAsnTyrIleGlyLysAlaLys 443
1275 CGCACCCGCGGAAAGAAATTTGCTGTGAACTACATTCATTCGCGGCTTCAA 1324
443 snProGluIleLeuSerTyrThrGluSerGlnLeuValGluValValAsp 459
1325 ACACCGGAATTCGTCCAAAGTCTGAAGGTGAGTTAGTGAAGCAGTTGAC 1374

460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVal 476
1375 AGAGATTTGAGGAAATGCTAATTAAGCTAATTAAGGACCGAATCCACTTAA 1424
476 IValGlyValArgValTyrProGlnAlaIleProGlnPheLeuValGly 493
1425 AATTAGAGTATGAGTATGGCCCTCAAGCCATTCCTCAGTTTGTAGTTGGTC 1474
493 ILeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
1475 ACTTTGATATTCCTTACACGCGCTAAATCATCTAACCTTCCGGGCTTAC 1524
510 GluGlyLeuPheLeuGlyLysAsnTyrValSerGlyValAlaLeuGlyArg 526
1525 GAAGGCTATTTTGGTGGCAATTAACCTGCTGTTAGCCCTTATGGCCG 1574
526 gCyValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 543
1575 GTGTGTAGAAAGCGCATATGAAACCGCGATTTGAGGTCAACAACCTCATGT 1624
543 erArgTyrAlaTyrLys 548
1625 CACGGTACCGCTTAAACAAG 1641
seq_name: gb_d12:ATHPPOX
seq_documentation_block: 1614 bp mRNA PIN 06-FEB-1999
LOCUS ATHPPOX Arabidopsis thaliana mRNA for protoporphyrinogen oxidase, complete cds.
DEFINITION Arabidopsis thaliana CDNA to mRNA.
ACCESSION D83139 GI:1183006
VERSION D83139.1
KEYWORDS protoporphyrinogen oxidase; POX.
SOURCE Arabidopsis thaliana CDNA to mRNA.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotids;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
s 1 (bases 1 to 1614)
REFERENCE
AUTHORS Narita, S.
TITLE Direct Submision
JOURNAL Submitted (17-JAN-1996) to the DDBJ/EMBL/GenBank databases.
AUTHORS Shin-Ichiro Narita, Kyoto University, Department of Biophysics,
Faculty of Science, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto,
Kyoto 606-01, Japan. E-mail:narita@molbio.dtiophys.kyoto-u.ac.jp,
Tel:075-753-4201, Fax:075-791-0271)
2 (bases 1 to 1614)
REFERENCE
AUTHORS Narita, S., Tanaka, R., Ito, T., Okada, K., Taketani, S. and Inokuchi, H.
TITLE Molecular cloning and characterization of a cDNA that encodes
protoporphyrinogen oxidase of Arabidopsis thaliana
JOURNAL Gene 182 (1-2), 169-175 (1996)
MEDLINE 97136707
FEATURES
source location/Qualifiers
1..1614
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
1..1614
/codon_start=1
/product="protoporphyrinogen oxidase"
/protein_id="BA11820.1"
/db_xref="GI:1877018"
/translacion="MELSLLRPTTOSLILPFSKRNRLNLYKPLRLRCVAVGGPTVGS
SKIEGGGTTITPCVIVGGIISGLCIAOLAKRHDAENLIVTFAKDKLVGNGNITLR
ENGFIMEBEENSPSPDPMITWVVDGSLDLDLVLPPTAPRVLLNNGKRLRVPYSKLT
DLPFDLMSIGKTRAGFALGTRPSPEEESVEEVRNLDDEVPERLIEEFCGTV
YAGPDKLSKAARAGVAKVKEONSGSITGTFALDERKNAAPAEADPRLPRCOGTV
GSPFRKGLRMLPEALISARLGSVKSLSWLSGITYKIESGIVLHTETPDGLVSVSKSV
MTPVSHAASGILRLYSANALSKLYPVAVVISYPEALIRTCCLIDGELKGFQD
LHPRTQGVETLGTIYSSSLPNNRAPPRRIILLNVIIGSSTWTGILSKSEGEFLVAVD
LRKMLIKPNSNDPLKIGVRVPOAIPQVIGHDILDITAKSSILTSSEGEFLVGNVY
AGVALGRCEVAGFEALIEVNNFMSRYAK"

BASE COUNT 416 a 334 c 421 g 443 t

ORIGIN

alignment_scores: Quality: 2075.00 Length: 536
Ratio: 4.235 Gaps: 4
Percent Similarity: 91.418 Percent Identity: 75.373

alignment_block: US-09-508-418-2 x ATNHPPGX
Align seg 1/1 to: ATNHPPGX from: 1 to: 1614

22 LeuAlaIleuAlaAsnArgThrSerPheIleProPheSerSerIleSerly 38
7 TTTATCTTCTTCGTCGCGACGACCTCAATGCGCTTCCGCTGTTTCGAA 56
38 sArTgAsnSerValAsnCysAsnGlyTirP.....ArgThrArgCysS 52
57 GCCCAAT...CTCCGATTAAGTGTATTADAGCCCTTAgACTCCGTTGTT 103
52 eYValAlaIleuAspIyrThrValProSerSerAlaValAspGlyIlePro 68
104 CAGTGGCCGGTGGACCAACCGTCCGATCTTCAAAAATCGAAGCGGAGGA 153
69 AlaAlaIleuLeu.....AspCysValIleValIleGlyIleSerG1 83
154 GGCACCCACCATCAGCAGGATTTGTGATTTGCGGGCGAGTATTAGTGG 203
83 yLeuCyIleAlaIleuValMetSerAlaAsnIyr.....ProA 96
204 TCTTTCATTCGCTCAGGCGCTTGTCTAAGCATCCGATGCTGCTCGGA 253
96 sIleuMetValThrGluAlaArgAspArgIleGlyIleAsnIleThrThr 112
254 ATTTATTTGTGAOCCGAGGCTTAGGAMTCGTGGAGGCAACATTAACACT 303
113 ValGluArgAspGlyTyrLeuIleuTrpGluGluGlyProAsnSerPheGlnPr 129
304 CGTGAAGAAATGTTTCTCTGCGGAAAGGTTCCCAATAGTTCCTCAACC 353
129 oSerAspPrometIleuThrMetAlaValAspCysGlyLeuIleuYAspAspL 146
354 GTCGTGATCTAATGCTCACTATGATGGTGAATAGTGGTAAAGATGATTT 403
146 euValIleuGlyAspProAsnAlaProArgPheValIleuTrpIlySgIlyS 162
404 TGGTGTGGGAGATCTACTGGCCCAAGTTTGTGTGGGAATGGGAAA 453
163 LeuArgProValProSerIleuThrAspLeuProPhePheAspIleuWk 179
454 TTGAGGCCGTTCCAMCGAAGCTAACAGACTTACCCTTTGATTTGAT 503
179 tSerIleProGlyIleuArgAlaGlyPheGlyProIleGlyLeuAArgP 196
504 GAGTATTTGGTGGAAAGATTAGAGCTGGTTTTGGTCCACTTGGCATTTCGAC 553
196 roSerProProGlyIleuSgIleuSerValGluIleuPheValIleuArgArgAsn 212
554 CGTCAACCTCCAGTCCGTTGAAGAATCTGTGGAGGAGTTTGTACGGGGTAA 603
213 LeuGlyGlyGluValPheGluIleuIleuIleuIleuIleuIleuIleuIleu 229
604 CTCGGGATGAGGTTTTGAGCGCCCTGATTTGAACCCTTTTGTTCAGGGT 653
229 IlyrValGlyAspProSerIleuSerIleuSerMetIleuAlaIleuPheGlyIleu 246
654 TTTATGCTGGTGTATCTTCAAAAATGATGATGATGATGATGATGATGATGAT 703
246 aITrPlyIleuGluIleuThrGlyIleuIleuIleuIleuIleuIleuIleu 262
704 TTTGGAAAATAGCAAAAATGGTGGAAAGCATTAATAGTGTGCTTTTAAAG 753

263 AlaIleuGlyIleuArgSerSerThrProIleuAlaProArgAspProArgle 279
754 GCAATTTCCAGGAGGAAAACCGCTCCAGGACAGACAGAAAGCCCGCCCT 803
279 uProIlySerProGlyIleuThrValIleuSerPheArgIlyIleuArgm 296
804 GCCAAAACACAGGCGCCAAACAGTGTCTTCTTCAAGAGGAGGACTTGGAA 853
296 etLeuProAspAlaIleuSerAlaArgIleuGlySerIleuIleuSer 312
854 TGTTCGCCAGAACAAATATTCGCAAGATTAGTACCAAAAGTTAAGTTGCT 903
313 TrPlyIleuSerSerIleuThrIleuSerGlyIleuGlyIleuIleuIleu 329
904 TGGAAAGCTCTCAGGATTCACATTAAGCTGAGAGCGAGAAATCAACTTAAC 953
329 rTYrGluIleuProGluIleuValValSerIleuIleuSerIleuValm 346
954 ATATGAGACTCCAGATGGTGTAGTTCCGTCAGAGCAAAAGAGTGTGTAA 1003
346 eThrValProSerTYrValAlaIleuSerAsnIleuIleuArgProIleuSerVal 362
1004 TGACGGTCCCATTCATGTTGCCAAGTGTGCTTCCGCGCCCTTTCGAA 1053
363 AlaAlaIleuAspAlaIleuSerAsnPheTYrTYrProValGlyAlaIleu 379
1054 TCTGCTGCAAAATGCACTCCAAACTTAAATTAACCCACAGTGGAGGAGT 1103
379 IThrIleuSerTYrProGluIleuValIleuArgAspIleuValAsp 396
1104 ATCTATCTCCGATCCCGAAGAGCAATCCGACAGAAATGTTTGTATGATG 1153
396 IyGluIleuIleuGlyPheGlyIleuIleuIleuIleuIleuIleuIleu 412
1154 GTAACATTAAGGTTTGGCAATTTGCATCCACGACGCAAGAGGTTGAA 1203
413 ThrIleuGlyThrIleuTYrSerSerIleuPheProAsnArgIleuPly 429
1204 ACAATTAGAACTATCACTACGCTCCCTCACTTCCAAATAGGCGGACCCGC 1253
429 sGlyArgValIleuIleuAsnTYrIleuGlyValIleuAsnProGlu 446
1254 CGAAGAAATTTGCTGTTGAACATTAATTTGGCGGCTTCAAAACCCGGAA 1303
446 IeIleuSerIlyThrGluSerGlnIleuValGluValIleuAspArgAspLeu 462
1304 TTTCTGCCAAGCTCAGAGTGTAGTGGAGAGCAGTTGACAGAGATTTG 1353
463 ArgIlyMetLeuIleuIleuProIleuAlaGlnAspProLeuValIleuGly 479
1354 AGGAAAATGCTAATTAAGCCCTAATTCAGCCGATCCACTTAATTAAGGAGT 1403
479 IArgValIleuProGluIleuIleuIleuIleuIleuIleuIleuIleuIleu 496
1404 TAGGATAGGCTCAGCCATTCCTCAGTTCCTAAGTGGTGGTCACTTTGATA 1453
496 hrIleuSerThrAlaIleuAlaIleuMetAsnAspAsnGlyIleuGlyIleu 512
1454 TCCTTGACACCGCTTAATTCATCTCAACGCTTCCGGGCTTACGAAAGGCTTA 1503
513 PheIleuGlyIleuAsnTYrValSerGlyValAlaIleuGlyIleuArgCysVal 529
1504 TTTTGGTGGCAATTAACGTCGCTGTTGAGCCCTTAGCGGGGTTGATAGA 1553
529 uGlyValIleuGlyIleuValIleuSerGluValIleuIleuIleuIleuIleu 546
1554 AGGCGCATATGAAACCGCGATGAGTCAACAACACTTCAATGTCACGGTACG 1603
546 IArTYrIleuS 548
1604 CTTACAAAG 1611
seq_name: gb_pat2:AX084742

seq documentation block: 1847 bp DNA PAT 09-MAR-2001

LOCUS AX084742 1847 bp DNA PAT 09-MAR-2001
DEFINITION Sequence 11 from Patent WO0112825.
ACCESSION AX084742
VERSION AX084742.1 GI:13274995

KEYWORDS soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I;
Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE Johnson, M. A., Volrath, S. L., Helfetz, P. B. and Law, M. D.
1 (bases 1 to 1847)
Herbicide-tolerant protoporphyrinogen oxidase
Patent: WO 0112825-A 11 22-FEB-2001;
Journal Syngenta Participations AG (CH)
FEATURES Location/Qualifiers
1..1847
/organism="Glycine max"
/db_xref="taxon:3847"
55..1686
/note="soybean protox-1"
/codon_start=1
/protein_id="CAC34013.1"
/db_xref="GI:13274995"
/translation="MWSVPEELIPPNQTLRPSLSHSPSEFTSPRRKPRRPRPRLI
RGSIAESTASPPKTRDASAPVDCVYVGGVSGICLAQALATRHANAVVVEARDRVG
GNTTTRRDYGLWDEGPNSEFQSDPMLTMVSDGLKDELVDLDPAPRFLVLRRLIE
VGRKLDLDFPDLFMEISGKIRAGFGLTRPPHESEVFEVFRNLGDEVEFRLLIE
PFGSGVAGDPSKLSMKAFKWKLEKNGSIIGTPEAIGERRGASKPPEDPRLPK
PGQVYGFSEFRKGLMPLDPAISARLGNKVKLSIKLDSGEVSLAYEPEGVVSL
OCKTVLRTIPSVASLILRPLSAAADALSKRYVPAVAVSTSYKPAIARSPCLIDGE
LKGFGDLHRSQGVETLGTIYSSSLEFPRRAPEGRIILNLTGCAINTGILSTDESLY
EYVRDLKILINPNAQDPFVGVRLMPQALIQFLVGHLDLDDVAKASIRNTGFEGLF
LGNVYSGVALRCVGEAYVAEYVNDPLTNRVYK"

BASE COUNT 473 a 416 c 453 g 505 t
ORIGIN

alignment_scores:
Quality: 2062.50 Length: 550
Ratio: 4.218 Gaps: 7
Percent Similarity: 88.909 Percent Identity: 73.273
alignment_block:
US-09-508-418-2 x AX084742 ..

Align seg 1/1 to: AX084742 from: 1 to: 1847
3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSer 19
102 TCTTCGCCCTCCTCCATTCGCCAACCCTTTCTCTCCATCCTCCACATC 151
19 TserProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSer 36
152 GAA.....ATTCC.....T 162
36 Le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
163 GCCTCTCCCTAACCCCTAAT.....CTACCTGCTC 194
52 ValAlaLysAspLysPheVal...ProSerSerAlaValAspLysGlyP 68
195 CATTCGGAGGAAATCCACCCGGCTCCGCCAAACAGAGACTCTCCGCC 244
68 roAlaAlaGluLeuAspCysValIleValAlaGlyAlaGlyIleSerGlyLeu 84
245 CC.....GTGACTGGCTGCGHGGGGAGGGGCTGAGGGGCTC 285
85 CysIleAlaGlnValMetSerAlaAsnTyr.....ProAsnLeuMetVa 99
286 TGCAATGCCCGAGGCTCCGCCAACACAGCCCAATGCAATGCAATGCTGCT 335

99 IthrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArg 116
336 CAGGAGGCCCGACGACCGGCTGGCGGCAACATCCACAGAGGAGGAGG 385
116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
386 ACGGATACCCTCTGGGAAAGAGGCCCAACAGCTTCCAGGCTCTGATCCA 435
133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuG 149
436 ATCCTCACCATGTGTGGAGAGTGTAAAGAMGAMGACCTGTTTGGGG 485
149 YAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgPro 166
486 GGAATCCTGATGACCTCGGTTGTGTGTGGAAACAGAGATGGAGGCCGG 535
166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
536 TCCCGCGGAAAGCTGACTGATTTGCCCTTCTTGTGACTGATGAGCATTTGGT 585
183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
586 GGCMAAATCAGGCGCTGCTGGCGCTGGAAATTCGGCCCTCCCTCC 635
199 OGlyHISGluGluSerValGluGluPheValArgArgAsnLeuGlyGly 216
636 AGGTCTTGAAGAGATTCGGTTGAAAGAGTTTCTGCTCCGAACTTGGTGTATG 685
216 lValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
686 AGGTTTGTGAAAGGATTAAGAGCCCTTGTGTCAGGGGTCTATGACAGGC 735
233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTyrLys 249
736 GATCCCTTCAAAATTAAGTAMGAAGCAGCAATTCGGAAAGTTTGGAGAGGT 785
249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
786 GGAATAAAATGCGTAGCATTATTTGTTGGAACCTTCAAGAGCAATCAAG 835
266 lArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
836 AGAGAAATGAGCCTCAAAACCCCTCGAGATCCCGGCTCCGCAAAACCA 885
283 LysGlyGluThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
886 AAAGGTGACGCTGTGATCTTCCGGAAAGGACHTTACCATGTGTGCTGA 935
299 alaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
936 TGCAATTTGCCGACGACTAGGCAACAAGTAAAGTATATGTTGGAAGCTTT 985
316 eIserIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
986 CAAGTATTAAGTAACTGAGTAGGAGAGAGTTCGACATATGAAACA 1035
333 ProGluGlyValValSerLeuGlnSerLeuArgSerIleValMetThrValPr 349
1036 CCAAGAGAGAGTGTCTTTTCCAGTGCANAAACTGTTGCTTCAACCATTTCC 1085
349 oSerTyrValAlaLysSerAsnIleLeuArgProLeuSerValAlaAlaAla 366
1086 TTCCTATGTTGCTAGTACATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
366 spAlaLeuSerAsnPheTyrTyrProProValAlaValAlaThrIleSer 382
1136 ATFCACATTTCAAAAGTTTATTTTACCCCTCCAGTTGCTGCAADTTTCCAMTACC 1185
383 TyrProGluGluAlaIleArgAspGluArgLeuValAspGlyGlyLeuLys 399
1186 TATCCAAAAGAGACTTATGATCAGAAATGCTTGTATGATGAGTGTGAGTGTGAA 1235


```

399  sglYpHeGlyGlnLeuHisProArGThrcInGlyValGluThrLeuGlyT 416
|||||
1236  GGGGTTGGTCAATTCACATCCAGCAGTACCGAAGAGTGGAAACATTAAGGAA 1285
416  hrlleYrSerSerleuPheProAsnArgAlaProLysGlyArgVal 432
|||||
1286  CTRRTRACAGCTCATTACCTATTCCTCCCAACCGAGCACCTCGAAGGGGTT 1335
433  LeuLeuLeuAsnTyrTlleglyLysAlaLysAsnProGluIleLeuSerly 449
|||||
1336  CTACCTCTGAATTAACCTTGGAGGAGAACCTAATTAAGTAAATTTTACGAA 1385
449  sThrcGlySerGlnLeuValGluValValAspArgAspLeuAArgLysMetL 466
|||||
1386  GACGGGACAGTGAACCTTGTGAAACAGTTGATGACGATTTGGAGAAATCC 1435
466  euIleYsProLysAlaGlnAspProLeuValValGlyValArgValTrrP 482
|||||
1436  TTRATAAACCAAAATGCGCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485
483  ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
|||||
1486  CCTCAAGCATTTCCACAGTTCCTTAGTGGCCATCTTGTATGATCTTAGATCT 1535
499  rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
|||||
1536  TCCTAAAGCTCTATACGAATAATCTGGGTTTGAAGGGGCTCTTCTTGGGGG 1585
516  lYAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGluLysAlaTyr 532
|||||
1586  GTAATTTATGTCTGTGTGTGGTGGGACGATGGCTTGAAGGAGACCCATAT 1635
533  GluValAlaSerGlnValThrGlyPheLeuSerArgTyrAlaTyrTllys 548
|||||
1636  GAGGTGACAGCTGAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1683
seq name: qb_pat2:AX084750

```

```

seq_documentation_block:
LOCUS       AX084750             1784 bp            DNA             09-MAR-2001
DEFINITION  Sequence 19 from Patent WO0112825.
ACCESSION   AX084750
VERSION     AX084750.1   GI:13275003
KEYWORDS
SOURCE      .
ORGANISM    Brassica napus
             Rapae.
             Brassica napus
             Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
             Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
             Brassicales; Brassicaceae; Brassica.
REFERENCE   1 (bases 1 to 1784)
AUTHORS    Johnson,M.A., Volrath,S.L., Heifetz,P.B. and Law,M.D.
TITLE      Herbicide-tolerant protoporphyrinogen oxidase
JOURNAL    Patent: WO 0112825-A 19 22-FEB-2001;
           Location/Qualifiers
FEATURES
   source
     /organism="Brassica napus"
     /db_xref="taxon:3708"
     47..1657
     /note="Ollseed rape protox-1 coding sequence"
     /codon_start=1
     /protein_id="CAC34016.1"
     /db_xref="GI:13275004"
     /translation="MDLSLRPQPLSPFSNPPFRSRPYKPLNLRCSVSGSVYSSST
     IEEGGGKRYVADCVIVGGTSGICTAOALVTRKPDAAKAVMVEAKADRVGNTIRPE
     EGGFLMEGRNSRPPSPMLTMVYDSSGLKDDLVIGDPTAARFVILRNKRLKPPVSKLFD
     LPFDLMSIGKIRAGFGALGIRSPGPRESEVFRKRNLDGDFERLLLEPCSGVY
     AGDPAKLSMKAAPFKVWKLEBNGSIIIGAFKAIOANKAPKTRDRPLRPGQTVG
     SFRKGLMPEPAISARLGDVKYKWLKLSITKLASGYSLETETPEGIVVOSKSYM
     TVPSHVASLLRPLSDSAAEALSKLYLPPVAAYISYAKFAIRSECLIDELGEGGLD
     HPTQKVEITGTTIVSSLPFRABPGRVILLINTYGAATNGIISKSRGELVEAVDRDL
     RKNLTKRSSTDPVLVGVKLPQALPQLIGHIDLVDAKAKSLSSSGHEGLFLGNVYA
     GVALRCKVEAGVETATOVNDPMSRIRAYK"

```

```

BASE COUNT      464 a      379 c      478 g      463 t
ORIGIN
alignment_scores:
  Quality: 2006.50
  Ratio: 4.189
  Percent Similarity: 89.533
  Percent Identity: 73.084
alignment_block:
US-09-508-418-2 x AX084750
Align seg 1/1 to: AX084750 from: 1 to: 1784
21  ProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerIleSe 37
   |||
   |||
   |||
68  CCGGACGCATTCCTATGCG.....CCATTTCCAAATCCATT 102
37  rLysArgAsnSerValAsnCysAsnGlyTTPArgThrArgCysSerValA 54
   |||
   |||
   |||
103 TCCTCCGG...TCGGGTCCTTCAAGCCTTCACACCTCCCGGTCCTCCGTAT 149
54  lAlYsAspTyrThrValProSerSerAlaValAspGlyValProAlaAla 70
   ::
   ::
   ::
150 CCGGTGATCCGTCGCGCTCTTCTACATCGAAGGCGGAGGAGGAGGT 199
71  Glu.....LeuAspCysValIleValGlyAlaGlyIleSerGlyLe 84
   ::
   ::
   ::
200 AAACCCGACAGCGGAGACTGCGTGTATCTGCGCGGAGAGANTCAGCGGCT 249
84  uCysIleAlaGlnValMetSerAlaAsnTyrPro.....AsnL 97
   |||
   |||
   |||
250 GTGCATTCGCCAAGGCGCTCGACGAAAGCACCCAGACGCTGCCAAAGATG 299
97  eumetValThrGluAlaAlaArgAspArgAlaGlyLysAsnIleThrThrVal 113
   ::
   ::
   ::
300 TGATGTGTCGAGGAGCGAAGGACCGTGTGGAGGAAATATCATCACCGCA 349
114  GluArgAspGlyTyrLeuTrrPglGluGlyProAsnSerPheGlnProse 130
   |||
   |||
   |||
350 GAGGAGCAAGGTTTCTATGGGAAAGAGTCCCAATATGCTTTCACCGCGT 399
130  rAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuY 147
   |||
   |||
   |||
400 TGATCTTATGCTCACTATGATGTAGATAGTGTGGTAAAGATGATCTAG 449
147  allLeuGlyAspProAsnAlaProArgPheValIleuTrrPlysGlyLeu 163
   |||
   |||
   |||
450 TCTTGGGAGATCTTACCTCCGAGGTTTGTGTGTGAAAGGAGGAGCTG 499
164  ArgProValProSerLysLeuThrAspLeuProPhePheAspLeuMetSe 180
   |||
   |||
   |||
500 AGCGCGGTTCCGCTCAGACTTAACTGACCTTCTTTGACTTGGATGAGAG 549
180  rIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProS 197
   |||
   |||
   |||
550 TATTGGAGGAAAGATTAAGACTGGTTTGGTCCATTGATTCGACACTT 599
197  erProProGlyHisGlnGlnSerValGlnGlnPheValArgArgAspLeu 213
   |||
   |||
   |||
600 CACCTCCCGGTCGTAGGAATCAGTGAAGGTTTGTAAAGGCGGTAACCTT 649
214  GlyGlyGluValPheGluArgLeuIleGluIuProPheCysSerGlyValTyr 230
   |||
   |||
   |||
650 GGTGATGAGGTTTTTGAACCGCTTATTAACCCCTTTTGTCTAGGTGTTTA 699
230  rValGlyAspProSerLysLeuSerMetLysAlaAlaPheGlyLysValT 247
   |||
   |||
   |||
700 TGCGGGAATCCTCGTGAACCTGAGATCAAAAGCAGCTTTTGGGAAGGTTT 749
247  rPlysLeuGlnGluThrGlySerIleIleGlyLysIleThrPheLysAla 263
   |||
   |||
   |||
750 GAAAGCTAGAGGAAATGTTGGAGACATCTTGGTGTCTTTTAAAGGCA 799

```

```

264 1lelygluargserSerThrProlysalabProargpProargLeuP 280
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
800 AFTCAAGGAAATAAAAGCTCCAGACAAACCCGAGATCCGGCTCTGCC 849
280 oLysProlysglyInthrValGlySerPheargLysGlyLeuargMetL 297
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
850 AAAGCCAAAGGCCCAACTGTTGGTTCCTGAGGAAAGACTCCAAATGC 899
297 euProspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerThr 313
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
900 TGCCAGAGCAATTCCTCCGAAAGTGGGTGACAAAGTGAAGTTCCTGG 949
314 LysLeuSerSerIleThrLysSerGluLysGlyLysThrHisLeuThrTy 330
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
950 AAGCTCTCAAGATACCTAACCTGGCCAGCCGAGAAATPACTTAACTTA 999
330 rGluThrProGluGlyValValSerLeuGlnSerArgSerIleValMetT 347
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1000 CGAAACTCCGGGGGTATATGACCTGTACAGACGCAAAAGTGTAGTGATGA 1049
347 hrValProSerTyrrValAlaSerAsnIleLeuArgProLeuSerValAla 363
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1050 CTGTGCCATCTCATGTCTGTAGTACTCTTGGCCGCTCTCTGTGATTC 1099
364 AlaAlaAspAlaLeuSerAsnPheTyrrProProValGlyAlaValThr 380
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1100 GCAGCTGAAGCGCTCTCAAAACTACTATCCGCCAGTGGAGCCGATTC 1149
380 rIleSerTyrrProGlnGluAlaIleArgAspGluArgLeuValAspGlyG 397
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1150 CATCTCAATACCGGAAAGCAATCCGAAAGCAATGCTTAATAGATGGTG 1199
397 LuLeuLysGlyPheGlyGlnLeuHisPProArgThrGlnGlyValGluThr 413
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1200 AACCTAAAAGGCTGGCCAGTTCGATCCACGACGCAAAAAGTGGAAACT 1249
414 LeuGlyThrIleTyrrSerSerLeuPheProAsnArgAlaProLysGly 430
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1250 CTTGGAAATATATACAGTTCATTCGCTTTCGCCAACGAGCACCCGCTGG 1299
430 yArgValLeuLeuLeuAsnTyrrIleGlyGlyAlaLysAsnProGluIleL 447
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1300 AAGATATATGCTATTTGAACCTACATCCGGTGGAGCTAACCAACTGGGATCT 1349
447 euSerLysThrGluSerGlnLeuValGluValValAspArgAspLeuArg 463
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1350 TATCAAGTCCGAAAGGTAGTATTAGTGGAAAGCATAGACTTTGAGCG 1399
464 LysMetLeuIleLysProLysAlaGlnAspProLeuValValGlyValAr 480
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1400 AAGATGCTGATAAAGCCAAAGCTCGACGATCCACTTGTACTTGGAGTAAA 1449
480 gValITrProGlnAlaIleProGlnPheLeuValGlyHisIleuAspThrL 497
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1450 ATTTAGGCTCAAGCCATTCCTCAGTTCGATGTAGTGCACACTTGTATTTGG 1499
497 euSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPhe 513
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1500 TAGACGCGCCAAAGCATTCGCTCTGTCATCTGTGCATGAGGGCTTATTC 1549
514 LeuGlyLysAsnTyrrValSerGlyValAlaLeuGlyArgCysValGluG 530
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1550 TTGGGTGGAAATTTACGTTGGCCGGTGTAGCANTGGGTGGTGTGGAGG 1599
530 yAlaITyrrGluValAlaSerGluValIThrGlyPheLeuSerArgTyrrAlar 547
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1600 TGCTTATGCAAACTGCAACCCCAAGTGAATGATTCATGTCAAGGTATCTT 1649
547 yLys 548
|||||
1650 ACAAG 1654

```

seq_name: gb_p12:AF160961

```

seq_documentation_block: 1925 bp mRNA PLN 01-OCT-1999
LOCUS AF160961
DEFINITION Cichorium intybus protoporphyrinogen IX oxidase (ppx1) mRNA,
complete cds; nuclear gene for plastid product.
ACCESSION AF160961
VERSION AF160961.1 GI:6002911
KEYWORDS
SOURCE
ORGANISM
Chicory.
Cichorium intybus
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
II; Asterales; Asteraceae; Lactuceae; Cichorium.
1 (bases 1 to 1925)
2 (bases 1 to 1925)

```

```

REFERENCE
1 Adomat,C. and Boeger,P.
Cloning, Sequence, Expression and Characterization of
Protoporphyrinogen IX Oxidase from Chicory
Unpublished
2 (bases 1 to 1925)
Adomat,C. and Boeger,P.
Direct Submission
Submitted (21-JUN-1999) Plant Biochemistry, University of Konstanz,
Konstanz
Location/Qualifiers
1. 1925
/organism="Cichorium intybus"
/db_xref="taxon:13427"
/feature="variety: foliosum"
1. 1925
/gene="ppx1"
91. 1758
/EC_number="1.3.3.4"
/codon_start=1
/product="protoporphyrinogen IX oxidase"
/protein_id="AF00194.1"
/db_xref="gi:6002912"
/translaction="MFSLEFDVCSLNCRCRSMSSLPPEYSGGSLRKNRPRLTYLTPSPAR
KNNRRFRCSIAKSDSPITPTISNERNFSDPLDVCYVAGISGICIAQALATKHAASV
DVTYEARDRRGGNISTVERDGYIWEGRSPSPSDMLTMVYDSDGLKDLVYIGDPTA
PRVYLVWGDLKPVPSKRPADLPEFDLMSFPEKLRGFEALGFRSPSPURESVEEVRK
NLGDEVERLLEPPCGSVYAGDPSEKLSMLAFKVMNLKEDNGSIVGGARKAIQDRKN
SOKPDPDRLPKPGQTVSFRKQOAMLPVALISRLSRYKLCMLKLSISKLENRGN
LTYEPQGESLQKTIYMTVPPVYVASDILRLPTISGLAADLISREYVPPVAASISYK
DAIRADRILISQILKRGQILHPRSGCVETLFTIYSSLPFRARAPRVLILNTGCAFN
PELISKTEGELVYDVPDRLRMLIRRAEPPRLTIGVYMPRAIPEPLIGHYDILDSAK
AALSSGGQUGFLGNGYSGVALKCVAAIVDVAAEVMPFLSOGVTK"
BASE COUNT 507 a 398 c 496 g 524 t
ORIGIN

```

```

alignment_scores:
Quality: 2004.50 Length: 536
Ratio: 4.194 Gaps: 3
Percent Similarity: 89.179 Percent Identity: 71.455
alignment_block:
US-09-508-418-2 x AF160961 ..
Align seg 1/1 to: AF160961 from: 1 to: 1925
18 SerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPhe 34
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
160 TCTGGTGGGCTCGTTGACCTCAAGATTCCTAGGTACCTAATACGTATAG 209
34 rSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpParThrArgC 51
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
210 TCCGGCCGCAATCGCAA.....TGCATATAGGTGGAGGCTCCCGCT 247
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
51 ySerValAlaLysAspTyrrVal...ProSerSerAlaValAspGly 66
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
248 GCTCTATAGCCAAAGATTCGCCCAATTTACTCTCCCATTTCAAAATGAGTTC 297

```

67 GYPRoAlAlaGluLeuAspCyValIleValIGIYAlaGlyIleSerGI 83
 : : : : :
 298 AACCTCCAGCCATGTTGGACTGTCATGTTGGCGCCGCATTAAGGG 347
 83 YLeuCYsIleAlaGluValMetSerAlaAsnTYR.....ProA 96
 : : : : :
 348 CCTTTGCATTTGGCGAGCCCTTACGCAATAACAGCCCTCCGCTCCGG 397
 96 sblEuMetValThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThr 112
 : : : : :
 398 ATGTGATTCGCACCCGAGCAGACAGAGAGTGGGGGTAAATATACAGC 447
 113 ValGluArgAspGlyTYrLeuTYrProGluGluProAsnSerPheGluPr 129
 : : : : :
 448 GTTGAAGGGATGGCTATCTGGGAAAGGTCTTAACAGCTTCCAGCC 497
 129 oSerAspProMetLeuThrMetAlaValAspCySglLeuIlyAspAspPL 146
 : : : : :
 498 ATCTGATGCCATGTCTCACCATGGTGGATAGTGGTGAAGGATGAT 547
 146 euValIleuGlyAspProAsnAlaProArgPheValLeuTYrPlySglYlyS 162
 : : : : :
 548 TGGTGTAGGGTGAACCCAGCACCCTTGTATATAGAGAGAGTGT 597
 163 LeuArgProValProSerIlyLeuThrAspLeuProPhePheAspLeu 179
 : : : : :
 598 TTGAACCGGTTCTCTTCCAAACCGGCTGACCTCTTCTTGGACCTCAT 647
 179 tSerIleProGlyIlyLeuArgAlaGlyPheGlyProIleGlyLeuAsp 196
 : : : : :
 648 GAGCTTTCGGAAAACTCAGACCCGCTTGGTCTGGATTCGCTC 697
 196 roSerProProGluIlySglIlySerValGluIlePheValArgArgAsn 212
 : : : : :
 698 CTTGACCTCCAGATCGGAAAGAAAGGTTGAGGATTTGTTAGACCTAAT 747
 213 LeuGlyGlyGluValIlePheGluArgLeuIleGluProPheCysSerGIyVA 229
 : : : : :
 748 CTTGGAGATGAGGATTTTCGAAACGCTTGTATGAAACCTTTTGTCCAGGTGT 797
 229 ITYrValGlyAspProSerIlyLeuSerMetIlySAlaAlaPheGlyIlyS 246
 : : : : :
 798 TTATGCTGCTGATCCAAACAACTTACTATGAAAGCAGCATTTGGGAAAG 847
 246 alTYrPlyLeuGluIlyThrGlyIlySerIleIleGlyIlyTYrPheIlyS 262
 : : : : :
 848 TCTGGAATCTGGAGCAAAATGGTGGTACATTTGTGGTGGAGCCTTCAAG 897
 263 AlaIleIlySgluArgSerSerThrProIlySAlaIleProArgPhe 279
 : : : : :
 898 GGTATTTACGAGACAGAAATGATGCAATTTGTGGTGGAGCCCGAGGTT 947
 279 uProIlySProIlySglIlyIlePheValGlySerPheArgIlyGlyLeuArg 286
 : : : : :
 948 ACCGAAACCAAGGGCCAAACCTGTTGATCTTTTAAAGAAAGGACAAAGCG 997
 286 etLeuProAspAlaIleSerAlaIleArgLeuGlySerIlySLeuIlySLeuSer 312
 : : : : :
 998 TGTTCCTAATGCAATCTCAACGAGGTTAGTACAGAGTGAATTTGTGT 1047
 313 TrPlyIlyLeuSerSerIleThrIlySerGlyIlySglYlyTYrHisLeuTh 329
 : : : : :
 1048 TCGAAGCTCACAGAGTATTTCAAAATTTGAGAAATAGAGATGATTAATTTGAC 1097
 329 rTYrIlyIlyThrProGluIlyValIleSerLeuGluSerArgSerIleIlyAl 346
 : : : : :
 1098 ATATGAAACCAACAGGATTTTGAAGTCTCAACATTAATAAATCTTCGTGA 1147
 346 ecThrValProSerTYrValAlaSerAsnIleLeuArgProLeuSerVal 362
 : : : : :
 1148 TGACTGTTCATCTAACGGGGAGTGACTGTGTCGCTCCGCTTTCGTTG 1197

363 AlaAlaAlaAspAlaLeuSerAsnPheTYrTYrProProValIGIYAlaVal 379
 : : : : :
 1198 GGTGACGAGATGCAATGTCAAAAATTTATATCCCTCCGTTGGACGCTGT 1247
 379 IThrIleSerTYrProGluIlyAlaIleArgAspGlyIlyrLeuValAspG 396
 : : : : :
 1248 ATCAATTTATATCCAAAGACCAATTCGTCATCCAGCGCTGATTTGATG 1297
 396 IyGluLeuIlySglIlyPheGlyIlyLeuHisProArgTYrGlnGlyValGlu 412
 : : : : :
 1298 GTCAACTCAAGGTTTGGGCAATTTGCATCCAGCAAGTCAAGGGGTGGAA 1347
 413 ThrLeuGlyThrIleTYrSerSerLeuPheProAsnArgAlaProIly 429
 : : : : :
 1348 ACTTGAAGTACAGATTCATCTCTTTTCCCTAACCCGAGCCGCCACC 1397
 429 sGIYAlaGlyValLeuLeuAsnTYrIleGlyIlyValAlaIlyAsnProGlu 446
 : : : : :
 1398 TGGAAAGGTTCTGCTTGAATACATTCGAGAGGCTTACAAATTCCTGAA 1447
 446 IeLeuSerIlyThrGluSerGluLeuValIleValValAspArgAspLeu 462
 : : : : :
 1448 TTCTATCAAAAGACGAGGGGAAATTTGTGATTCGGGTGAGCCGGAGCCTA 1497
 463 ArgIlyMetLeuIleIlySProIlySAlaGluAspProLeuValValGly 479
 : : : : :
 1498 CCGACGATTCCTGATTAAGCCGTGATGGGAAAGATCCATTAACCTTGGGG 1547
 479 IArgValTYrProGluAlaIleProGluIlePheLeuValGlyHisLeuAsp 496
 : : : : :
 1548 GCGGGTGTGGCTCCGACCAATCCCGGAGTTTGTGATCGGTCATTTWAGCA 1597
 496 hrLeuSerThrAlaIlySAlaIleMetAsnAspAsnGlyLeuGluIlyLeu 512
 : : : : :
 1598 TTCTAAGATTCGCAAAAGCTCTGTGATGACGGTGGATTCACAAAGTATG 1647
 513 PheLeuIlyGlyAsnTYrValIleSerGlyValAlaIleGlyIlyArgCysValG 529
 : : : : :
 1648 TTYCTGGTGGACACTATGTTGTGGTGGTGGCTTTAGTAANAATGTGCA 1697
 529 uGIYAlaTYrGlyIlyValAlaSerGlyIlyValIleThrGlyPheLeuSerArgTYr 546
 : : : : :
 1698 GCGCTTATGATGATGTTCCCGCTGAGTAAATGAACTTTTGTCCCAAGGG 1747
 546 IaTYrIlyS 548
 : : : : :
 1748 TGTACAAAG 1755

seq_name: gb_p12:AF218052

seq_documentation_block: 1608 bp mRNA PIN 15-FEB-2000
 LOCUS AF218052 Zea mays protoporphylinogen IX oxidase mRNA, complete cds.
 DEFINITION Zea mays protoporphylinogen IX oxidase mRNA, complete cds.
 ACCESSION AF218052
 VERSION AF218052.1 GI:6715440
 KEYWORDS
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 1608)
 Volrath, S.L.
 DIRECT SUBMISSION
 Submitted (20-DEC-1999) Novartis Agricultural Biotechnology, 3054
 Cornwallis Road, Research Triangle Park, NC 27709, USA
 LOCATION/Qualifiers
 1. 1608
 /organism="Zea mays"
 /cultivar="B73 Inbred"
 /db_xref="taxon:4577"
 1. 1608
 /codon_start=1

CDS

```

/product="protoporphyriinogen IX oxidase"
/protein_id="AAF26417.1"
/db_xref="GI:6715441"
/translation="NWAAARTRTAMATRAASPLNGLTRIPARLRHRRLRGLSVRCAAVAGAAE
AFASSTARISADCVVYVGGISGLCTRAQALATRHGVDLVLEARARPEGNITTYERE
EGYLWEEGPNRSFQSDPVLVMAVDGLKDDLVFVPMAPRVLWEGKLRVPSKPADL
PEFDLMSIGKLRAGLALGIRVPPGREGVEEVRNRNLAEVEERLIEPSCGYVA
GDPDKLSMKAARFQKWLRETEGSLIIGTIRKIQSRNKPVRDARLPPKSGQTVS
FRKGLMLNPAITSSISGSKVLSMKLTSITKSDDDGYLVEVETPGVYVQAKGQVIMT
IPSYVSNLIRPLRSDPADALSRFYPVPAVAATVYSPKFAIRKECLIDGELQGRGDLR
PRSGVETLGTIYSSLSLFPNRA PDGRVYLLNLTGGATMTGIVSRKESPLVAVDRDR
KMLINSTAVDPLVLCYRVPQAIPOPVYGHDDLLAAAKALDRGSDGLFLGNNYVAG
VALRCKVEGYESASQISDFLTKYAK"
BASE COUNT 360 a 430 c 464 g 354 t
ORIGIN

```

```

alignment_scores:
Quality: 1989.50 Length: 545
Ratio: 4.260 Gaps: 6
Percent Similarity: 85.688 Percent Identity: 72.294

```

```

alignment_block:
us-09-508-418-2 x AF218052

```

```

Align seg 1/1 to: AF218052 from: 1 to: 1608

```

```

17 SerSerSerSerProLeuAlaIAlaPheLeuAsnArGThrSerPheIleProPh 33
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
31 ACCGCTGCATCCCGCCCTA.....CTCAACGGGACCCGA...ATACCTGC 71
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
33 eSerSerIleSerLysArgAsnSerValAsnCyAsnGlyTTrpArgThra 50
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
72 GCGGCTCCGCCATCCAGAGACACGCGCTCC..... 105
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
50 rGcYSerValAlaLysAspTyrThraVAlProSerSerAlaValAspGly 66
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
106 .....GCTGCTGTGGCGGCC 120
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
67 GLYProAla.....AlaGluLeuAspCy 74
||| ||| |:::|:::|:::|:::|:::|:::|
121 GCGGCGCCGAGGACCGGCATCCAGCGCGCGGCGCTGTCCGGGAGACTG 170
||| ||| |:::|:::|:::|:::|:::|:::|
74 sValIleValGlyAlaGlyIleSerGlyLeuCyIleAlaGlnValMetS 91
||| ||| |:::|:::|:::|:::|:::|:::|
171 CPTCGTGGGGGCGGAGCATCAGTGTGCGCTGTGACCGCGCAGCGGCGCTGG 220
||| ||| |:::|:::|:::|:::|:::|:::|
91 eAlaAsnTyr....ProAsnLeuMetValThnGluAlaArgAspArg 105
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
221 CCAGCGGCGACGGCTCGGGGACGTGCTTGTCAAGAGGCCCCCGCCCGC 270
||| ||| |:::|:::|:::|:::|:::|:::|
106 AlaGlyGlyAsnIleThrThrValGlnArg.....AspGlyTyrLeuTr 120
||| ||| |:::|:::|:::|:::|:::|:::|
271 CCGGCGGGAACATTAACACCAGTGTGAGCGCCCGAGGAAAGGGTACCTGTG 320
||| ||| |:::|:::|:::|:::|:::|:::|
120 pGluGluGlyProAsnSerPheGlnProSerAspProMetLeuThiMetA 137
||| ||| |:::|:::|:::|:::|:::|:::|
321 GGAAGGGGTTCCACACAGCTTCCAGCCCTCCGACCCGCTTCCACCAATGG 370
||| ||| |:::|:::|:::|:::|:::|:::|
137 LaValAspCysGlyLeuLysAspAspLeuValIleGlyAspProAsnAla 153
||| ||| |:::|:::|:::|:::|:::|:::|
371 CCGTGGACAGCGGACTGAAGGATGACTGTGGTTTGGGGACCCCAACCGCG 420
||| ||| |:::|:::|:::|:::|:::|:::|
154 ProArgPheValLeuTrpLysGlyLysLeuArgProValProSerLysLe 170
||| ||| |:::|:::|:::|:::|:::|:::|
421 CCGCGTTCCTGCTGTGGAGGAAAGCTGAGGCCGCTGATCCAAAGCC 470
||| ||| |:::|:::|:::|:::|:::|:::|
170 uThraspLeuProPhePheAspLeuMetSerIleProGlyLysLeuArg 187
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
471 GCGGACCCCGCTTCCGATTCATGACATCCAGGGAAAGCTCAGGG 520
||| ||| |:::|:::|:::|:::|:::|:::|
187 laGlyPheGlyProIleGlyLeuArgProSerProProGlyHisGlu 203
||| ||| |:::|:::|:::|:::|:::|:::|

```

```

521 CCGGCTGAGCGCGCTTGGCATCCGCCGCCCTCCAGCGCCGAGAAAG 570
204 SerValGluGlnPheValArgArgAsnLeuGlyGlyGluValPheGlu 220
||| ||| |:::|:::|:::|:::|:::|:::|
571 TCAGTGGAGGATTCGTGCGCCGACACCTCGGTGCTGAGGCTTTTGAGCC 620
||| ||| |:::|:::|:::|:::|:::|:::|
220 gLeuIleGluProPheCysSerGlyValIleValGlyAspProSerLysL 237
||| ||| |:::|:::|:::|:::|:::|:::|
621 CCTCATTTAGGCTTTCGCTCAGGCTGTCTATAGCTGATGCACTTCAGAC 670
||| ||| |:::|:::|:::|:::|:::|:::|
237 eSerMetLysAlaIAlaPheGlyLysValIlePlyLysLeuGluThrGly 253
||| ||| |:::|:::|:::|:::|:::|:::|
671 TCAGCATAGAGGCTGCATTTGGAAAGCTTTGGCGGTTGGAAAGAACTGGA 720
||| ||| |:::|:::|:::|:::|:::|:::|
254 GlySerIleIleGlyGlyThrPheLysAlaIleLysGluArgSerSerTh 270
||| ||| |:::|:::|:::|:::|:::|:::|
721 GSTAGTATTATTGGTGCACATCAGAACAAATTCAGAGAGAGAGAAAGAA 770
||| ||| |:::|:::|:::|:::|:::|:::|
270 rProLysAlaProArgAspProArgLeuProLysProLysGlyGluThrV 287
||| ||| |:::|:::|:::|:::|:::|:::|
771 TCCAAAACCACCGAGGATWGCCCGCTCCGAAAGCCAAAGGGGAGAGCAG 820
||| ||| |:::|:::|:::|:::|:::|:::|
287 aGlySerPheArgLysGlyLeuArgMetLeuProAspAlaIleSerAla 303
||| ||| |:::|:::|:::|:::|:::|:::|
821 TTGCATCTTTCAGGAGAGGCTTGTCCATAGCTTCCAAATGCAATTAACATCC 870
||| ||| |:::|:::|:::|:::|:::|:::|
304 ArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSerIleThrLy 320
||| ||| |:::|:::|:::|:::|:::|:::|
871 AGCTTGGTGTAGTAAAGTCAAACTATGAACTCAGAGCACTTAAACAA 920
||| ||| |:::|:::|:::|:::|:::|:::|
320 sSerGluLysGlyGlyTyrHisIleuThrTyrGluThrProGluGlyVal 337
||| ||| |:::|:::|:::|:::|:::|:::|
921 ATCAGATGACAAAGGATATGTTTGGAGTATGAAAGCCCAAGAGGGGTTGG 970
||| ||| |:::|:::|:::|:::|:::|:::|
337 aSerLeuGlnSerArgSerIleValIleThrValProSerTyrValAla 353
||| ||| |:::|:::|:::|:::|:::|:::|
971 TTTCCGTCAGGCTTAAAGATGTTATCATGACTATTCATATATATGTTGCT 1020
||| ||| |:::|:::|:::|:::|:::|:::|
354 SerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAlaLeuSerAs 370
||| ||| |:::|:::|:::|:::|:::|:::|
1021 AGCAACATTTTGGCGTCCACCTTCAAGGATGCTGCACATGCTCTATCAAC 1070
||| ||| |:::|:::|:::|:::|:::|:::|
370 nPheTyrTyrProProValGlyAlaValAlaThrIleSerTyrProGlnGlu 387
||| ||| |:::|:::|:::|:::|:::|:::|
1071 ATTCATATATCCACCGCTGCTGTGTAACGTTTCTATVCAAAAGGAAAG 1120
||| ||| |:::|:::|:::|:::|:::|:::|
387 laIleArgAspGluArgLeuValAspGlyGluLeuLysGlyIlePheGlyGln 403
||| ||| |:::|:::|:::|:::|:::|:::|
1121 CAATTAAGAAAGAAATGCTTATGATGGGAACTCCAGGGCTTTGGGCGAG 1170
||| ||| |:::|:::|:::|:::|:::|:::|
404 LeuHisProArgThrGlnGlyValGluThrLeuGlyThrIleTyrSerSe 420
||| ||| |:::|:::|:::|:::|:::|:::|
1171 TTGCATCCACGATGACAGAGGATGAGCATTAGGAACAATATATACATTC 1220
||| ||| |:::|:::|:::|:::|:::|:::|
420 rSerLeuPheProAsnArgAlaProLysGlyArgValIleuLeuLeuAsnT 437
||| ||| |:::|:::|:::|:::|:::|:::|
1221 CTCACCTCTTCCAAATCTGTCTCTGACAGGATGAGGCTTTCCTTAACCT 1270
||| ||| |:::|:::|:::|:::|:::|:::|
437 YrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThrGluSerGln 453
||| ||| |:::|:::|:::|:::|:::|:::|
1271 ACATATGAGAGGCTTACAAACACAGAGAAATTTGTTCCAAAGACTGAAAGGAG 1320
||| ||| |:::|:::|:::|:::|:::|:::|
454 LeuValGluValAlaAspArgAspLeuArgLysMetLeuIleLysProLy 470
||| ||| |:::|:::|:::|:::|:::|:::|
1321 CTGGTCCAAAGCAGTGTGACCCGTGACCTCCGAAAAATGCTTAAATTTTAC 1370
||| ||| |:::|:::|:::|:::|:::|:::|
470 saLagInAspProLeuValAlaGlyValArgValTrpProGlnAlaIleP 487
||| ||| |:::|:::|:::|:::|:::|:::|
1371 AGCAGTGAACCTTATGCTTGGTGTGAGTGTGGCCCAAGCCATAC 1420
||| ||| |:::|:::|:::|:::|:::|:::|
487 roGlnPheLeuValGlyHisLeuAspThrIleuSerThrAlaLysAlaAla 503
||| ||| |:::|:::|:::|:::|:::|:::|
1421 CTCAGTTCCTGTTAGGACATGTGATCTTGTGAAAGCCGAAAGACTGCC 1470
||| ||| |:::|:::|:::|:::|:::|:::|

```

504 MetAsnAspansnlyLeuGlueGlyLeuPheLeuGlyAsnTyValSe 520
 1471 CTGACCGAGTGGCTACACGATGCTGTTCCZTAGGGAGGAACTATGTGC 1520
 520 rGlyValAlaLeuGlyArgCysValGlyIleAlaTyrgIuValAlaSe 537
 1521 AGGAGTTCCTGCTGGCAATGCGTGAAGGCGGCTGATGAAGAGTGCCTGC 1570
 537 IuValThrglyPheLeuSerArgTyAlaTyrls 548
 1571 AATATACCTGACTTCTTGACCAAGTATGCTACAG 1605
 seq_name: gp_pat2:AX084740

seq_documentation_block:
 LOCUS AX084740 1811 bp DNA PAT 09-MAR-2001
 DEFINITION Sequence 9 from Patent W00112825.
 ACCESSION AX084740
 VERSION AX084740.1 GI:13274993
 KEYWORDS
 SOURCE
 ORGANISM
 'bread wheat.
 Triticum aestivum
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae;
 Triticum.
 1 (bases 1 to 1811)
 Johnson, M.A., Volrath, S.L., Heifetz, P.B. and Law, M.D.
 Herbicide-tolerant protoporphyrinogen oxidase
 Patent: WO 0112825-A 9 22-FEB-2001;
 Syngenta Participations AG (CH)

FEATURES
 source
 location/Qualifiers
 1..1811
 /organism="Triticum aestivum"
 /db_xref="taxon:4565"
 <3..1589
 /note=" wheat protox-1"
 /codon_start=1
 /protein_id="CAC34012.1"
 /db_xref="GI:13274994"
 /translation="ATMATATVAASPLRGRVTRPRHRVRCATASATETPAPCV
 RLSAEIVAGISLCTAQAALATRYGVSVDLVEARDRGGNTTVERDEGILME
 GPNSEFOPSDPVLTMVAVDGSKDDIVFGDPAAPRVTVEKLRVPSKRPGLPPSIS
 IPKLRAGLGAIGRRPPPRESEVEEYRNRNIGAEVFEELIEPCGYVAGDPSKIS
 MKAIFKRWRIEIEIGSIIIGSTIAIDDKRNPDPPLPAPKGTIVASFKGLAM
 LPAIARLSKVKLISWKLSITKADQGVGLGEPEEGLVSOAKSVIMTIFRSYVA
 DILRPLSIDADALSKFYPPVAIVVSYPKAIRKELIDGEGDQDHPSSQVIE
 TLGIYSSLEPNRAPAGRVILLNYIGSSTNTGIVSTESDLDGAVDRDLRKMILNR
 AADPLALGVRWVPAIROPFLIGHDRLAAKAKSALGGGVDGLFLGGVYVAGVALGRCI
 EGAYESASOVSDPTTKYAK"

BASE COUNT 423 a 463 c 517 g 408 t
 ORIGIN

alignment_scores:
 Quality: 1987.00 Length: 505
 Ratio: 4.310 Gaps: 3
 Percent Similarity: 91.287 Percent Identity: 75.050

alignment_block:
 US-09-508-418-2 x AX084740
 Align seg 1/1 to: AX084740 from: 1 to: 1811

48 ArgThrArgCysSerValAlaLysAspTyrrhrValProSerSerAlaVa 64
 78 CCGCCGCGTGGCGTACCGCGAGCGCGGACCGGACGCTCCGCGCGGCG... 125
 64 IaspglyglyProAlaAlaLysAspCysValIleValGlyAlaGlyT 81
 126 ...CCCGCGTGGCGTCCGCGGCAATGCGTCAATGTGGCGCGCGGCA 171
 81 IeSerGlyLeuCysIleAlaLysValMetSerAlaAsnTyrr.....Pro 95

172 TCAGCGGCGCTGCACCGCGAGGGCGTGGCCACCACCGATACGGCGGTAGC 221
 96 AsnLeuMetValThrGluAlaArgAspArgAlaGlyIleThrTh 112
 222 GACTGCTGTCACGAGGAGCCCGCGAGCCCGCGGCGCACAATCACCC 271
 112 rValGluArg.....AspGlyTyrlLeuTrpGluGlyIleProAsnSerP 127
 272 CGTCGAGCGTCCCGACGAGGGGTTACCTGTGGGAGGAGGACCAACAGCT 321
 127 heGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuLys 143
 322 TCCAGCCCTCCGACCCCGTCCCTCACCATGGCCGTCGACAGCGGCGCTCA 371
 144 AspaSpLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrpLys 160
 372 GATGACTTGGTGTGGGAGCCACCAAGCGCGCGCTTCGTGCTGTGGGA 421
 160 sGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePhe 177
 422 GGGGAAAGCTGAGCGCGGTGCCGTGGAAGCCAGCGGACCTGCTTCTTCA 471
 177 sPLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleGly 193
 472 GCCTCATGATGATCCCTGGGAAAGCTCAGCGCGCGCTTGGCGCGCTGC 521
 194 LeuArgProSerProProGlyHisGluGluSerValGluGlnPheValAr 210
 522 ATTCGCCACACTCTCCAGCGCGGAGGAGATCGGTGGAGAGTGTGGCG 571
 210 garGAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCysS 227
 572 CCGCAACCTCGGTGCGGAGGTCTTGAAGCGCGCTCATGAGAGCTTTCGT 621
 227 erGlyValThrValGlyAspProSerLysLeuSerMetLysAlaAlaPhe 243
 622 CAGGTGATGCTGGTGTATCTTCGAAAGCTTAGTATGAAAGCGCGATTT 671
 244 GlyLysValTrpLysLeuGluGluThrGlyLysSerIleIleGlyGlyTh 260
 672 GGGAAAGCTGAGAGTTGAGAGATTTGAGAGTATGATTTATGTTGGTGAAC 721
 260 rPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAsp 277
 722 CATCAAGCGATTCAGATTAAGGAAAGCAACCCCAACCGCGAGGATTC 771
 277 roArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysGly 293
 772 CCGGACTTCCCGCACCAAAAGGAGAGAGGTGATTCATTCAGGAAAGGTT 821
 294 LeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLys 310
 822 CTAGCATGCTCCCGAATGCCATTCGCAATTCAGTGGGTGTAAGCA 871
 310 sLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTh 327
 872 GCTGTCATGAGGAGGCTTACGACATTCACAAAGCGCGCAACCAAGATATG 921
 327 IsLeuThrTyrgLnuThrProGlnGlyValIleSerLeuGlnSerArgSer 343
 922 TATTAGGTTATGAACACACGAAAGGACTTGTTCAGTACAGCGGTAAAGAT 971
 344 IleValMetThrValProSerTyValAlaSerAsnIleLeuArgProLe 360
 972 GTTATCAATGACATCCCGTCAATGTTGTGATGATTCCTGGCCGCACT 1021
 360 uSerValAlaAlaLysAlaLeuSerAsnPheTyrrProProValG 377
 1022 TTCATATGATGACAGATGACCTCAAAATTCATTATTCAGCCAGATTTG 1071
 377 LyAlaValThrIleSerTyrrProGlnGluAlaIleArgAspGluArgLeu 393

```

1072 CTGCTGTAACCTGTTTCATATCCAAAAGAGCTATTAGAAAAGAAATGCTTA 1121
394 ValAspGlyGluLeuLeuLeuGlyPheGlyGlnLeuHisProArgThrGlnG1 410
1122 APTGATGGAGGAGCTCCAGGGGTTCGGCCAGTTGCACTCAGCTACCAAGG 1171
410 YValGlnThrLeuGlyThrIleYrSerSerSerLeuPheProAsnArg 427
1172 AGTGGAGACTTATAGGACAAATATAGCTCTCTCTCTCTCTCTCTCTCTCTCT 1221
427 IaProLysGlyArgValLeuLeuLeuLeuAsnYrIleGlyGlyAlaLysAsn 443
1222 CTCCTGCGTCAAGAGGTGTACTCTGTAACATATCCGGGGGTTCACAAAT 1271
444 ProGluIleuSerLysThrGluSerGlnLeuValGluValAlaAspArg 460
1272 ACAGGAAATGCTCCCAAGACTGAGAGGTGACTAGAGAGCCCTTGGACCG 1321
460 qAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuValY 477
1322 TGACCTCGAAAATAAGTTGATTAACCTAGAGCAKCAKCAKCAKCAKCAKCA 1371
477 a1GlyValArgValLrProGlnAlaIleProGlnPheLeuValGlyHis 493
1372 TAGGGGTTCGAGTGGCCACACAGCAATACCAAGTTTGTGATTTGGGCGAC 1421
494 LeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuG1 510
1422 CTTGATGCCCTTGTCTGCAAAATCTGCACCTGGGCCCAAGGGCTTACGA 1471
510 uGlyLeuPheLeuGlyGlyAsnYrValSerGlyValAlaLeuGlyArg 527
1472 CGGGTGTCTCTAGAGAACTACGCTGCGAGAGTTCCTTGGGGCCGAT 1521
527 ySValGluGlyAlaTrpGlyValAlaSerGluValThrGlyPheLeuSer 543
1522 GCATCGAAGGCGCTGACGAGGTGCTCCACAAATCTGACTCTTGTGACC 1571
544 ArgTrpAlaTrpLys 548
1572 AAGTATGCTTACAG 1586

```

```

seq_name: qb_pat2:AX084736
seq_documentation_block:
LOCUS AX084736 1691 bp DNA PAT 09-MAR-2001
DEFINITION Sequence 5 from Patent WO0112825.
ACCESSION AX084736
VERSION AX084736.1 GI:13274989
KEYWORDS
SOURCE
ORGANISM Zea mays.
zma mays
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1691)
JOURNAL Johnson, M.A., Volrath, S.L., Helfetz, P.B. and Law, M.D.
Herbicide-tolerant protoporphyrinogen oxidase
Patent: WO 0112825-A 5 22-FEB-2001;
Syngenta Participations AG (CH)
FEATURES
source
1.1691
/organism="Zea mays"
/db_xref="taxon:4577"
<1.1446
/note="Maize protox-1 c-DNA (not full-length)"
/codon_start=-1
/protein_id="CAC34010.1"
/db_xref="GI:13274990"
/translation="ADCVVWGGISGLCTAQAALATRHVGVYLVTEARARPGNITTY
ERREBXTLMEEGNSFQPSDPYITMAVDSGLKDDLVFGDPNAPRFLTMEGKTRPVPSK
PADLPEFDLWSTIGKLRAGLGIKRPPTGPPRESVEEYVKKRNLGAEVYERLIEPESQ
GYVAGDPSKISMKAARFKYWRLEETGSIIGTIKTIQERSKNPKPPRDARLKPKGQ

```

```

alignment_scores:
Quality: 1978.00 Length: 480
Ratio: 4.455 Gaps: 2
Percent Similarity: 92.500 Percent Identity: 78.750
alignment_block:
US-09-508-418-2 x AX084736 ..
Align seg 1/1 to: AX084736 from: 1 to: 1691
73 AspCyValIleValGlyValGlyIleSerGlyLeuCyIleAlaGlnA 89
|||||.....:|||||.....:|||||.....:|||||.....:
4 GACTGGGTGTTGGTGGCCGAGGCAATCACTGCTTGCACCCGCGCAGGC 53
89 ImeSerrAlaAsnTyr.....ProAsnLeuMetValThrGluAlaArgA 104
.....:.....:.....:.....:.....:.....:
54 GCTGGCCACGGCCGACGGCGGCGGCGGAGCGTCTGTCACGGAGCCCGCG 103
104 sPArgAlaGlyGlyAsnIleThrThrValGluArg.....AspGlyYr 118
|||.....:.....:.....:.....:.....:.....:
104 CCGCCCGCGCGGCAACATTACCCGCTGGAGGCCCGAGGAAAGGGTAC 153
119 LeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetLeuTh 135
|||||.....:|||||.....:|||||.....:|||||.....:
154 CTTCTGGAGAGAGGCTCCCAACAGCTTCCAGCCCTCCGACCCCGTCTCAC 203
135 rMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAspProA 152
|||||.....:|||||.....:|||||.....:.....:.....:
204 CATGGCCGTTGACAGCGGACTGAAAGGATGACTTGTGGGGGACCCAA 253
152 snAlaProArgPheValIleuTrpLysGlyLysLeuArgProValProSer 168
|||||.....:|||||.....:|||||.....:.....:.....:
254 ACGCGCGCGGCTTTCGTGCTGTGGAGGAGAGGAGCTGAGGCCCTGCCATCC 303
169 LysLeuThrAspLeuPProPhePheAspLeuMetSerIleProGlyLysLe 185
|||.....:.....:.....:.....:.....:.....:
304 AAGCCCGCGCACCTCCCGTTCCTTCGATTCAGATCAGAGCATCCAGGAAAGCT 353
185 uArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGlyHisG 202
|||||.....:|||||.....:|||||.....:.....:.....:
354 CAGGGCCGGCTTAGCGCGCTTGGCATCCGCCCTCTCCAGGCGCGG 403
202 IuGlnSerValGluGlnPheValArgArgAsnLeuGlyGluValAlpHe 218
|||||.....:|||||.....:|||||.....:.....:.....:
404 AAGAGTCAGTGGAGGAGATTCGTCGCCGCCGCACTCCGCTGAGAGTCTTT 453
219 GluArgLeuIleGluProPheCysSerGlyValYrValGlyAspProSe 235
|||||.....:|||||.....:|||||.....:.....:.....:
454 GAGCGCCTCATTTGAGCCCTTCTCTCAGGGTCTATGCTGATCATCTTTC 503
235 LysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuGluGlnT 252
|||||.....:|||||.....:|||||.....:.....:.....:
504 TAAAGCTCAGCATGAAAGGCTGCAATTTGGAAAGTTTGGGAAAGAA 553
252 hrcGlySerIleIleGlyGlyThrPheLysAlaIleLysGluArgSer 268
|||||.....:|||||.....:|||||.....:.....:.....:
554 CTGGAGGTAATATTATTTGGTAACCATCAAGCAATTCAGAGAGAGAGAGC 603
269 SerThrProLysAlaProArgAspProArgGluProLysProLysGlyG1 285
.....:.....:.....:.....:.....:.....:
604 AAGAAATCAAAACCACCGAGAGATGCCCTCCGAAACCCAAAAGGAGC 653
285 nThrValGlySerPheArgGlyGlyLeuArgMetLeuProAspAlaIleS 302
|||||.....:|||||.....:|||||.....:.....:.....:

```

654 GACAGTTGCATCTTTCAGGAGGCGCTTCCAGCAGCTTCCAAATGCCATTA 703
302 eraLaarLeuGlySerLyLeuLyLeuSerTrrpLyLeuSerSerIle 318
704 CATCCACCTGGGTAGTAAAGTCAAACTATCATGGAACACCCAGAGGATT 753
319 ThrLySerLeuLySerGlyTyrHisLeuThrTyrGluThrProGlu 335
754 ACAAAATCAATATGCAAGGATATGTTTGGAGTATGAAAGCCAGAGG 803
335 yValValSerLeuGlnSerArgSerIleValMetThrValProSerTyr 352
804 GGTGTTCGGGTGAGGCTAAAGTGTATCATGACTATTCCATCATAG 833
352 aLaLaSerAsnIleLeuArgProLeuSerValAlaAlaLaSerValLeu 368
854 TTGCTACCAACATTTTCCGCTCCACTTTCAGCGATGCTGCAGATGCTCA 903
369 SerAsnThrTyrTrpProProValGlyAlaValThrIleSerTyrProG 385
904 TCAGATTTATATTCACCCGGTGTGCTGTAAGCTGTTCCATCCAA 953
385 nGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLySerGlyPheG 402
954 GGAAGCAATTAGAAAAGAAATGCTTAATTTGATGGGAACTCCAGGCTTG 1003
402 LyGlnLeuHisProArgThrGlnGlyValAlaGluThrLeuGlyThrIleTyr 418
1004 GCCAGTTCATCCACCGTAGTCAAGGAGTTGAGACATTTAGAACATATATAC 1053
419 SerSerSerLeuPheProAsnArgAlaProLyGlyArgValLeuLeuLe 435
1054 AGTTCCTCACTCTTCCAAATCGTCTCCAGCGGTAAGGGGTACTTCT 1103
435 uAsnTyrIleGlyAlaValAsnProGluIleLeuSerTyrThrGluS 452
1104 AAACACTAGTAGGAGTGTACAAACACAGGAATTTTCCAAAGACGAA 1153
452 erGlnLeuValGluValValAspArgAspLeuArgLyMetLeuIleLys 468
1154 GTAGGCTGTGGAAGCAGTTGACCGCTCCGAAATAATGCTTATAAAT 1203
469 ProLyAlaGlnAspProLeuValValGlyValArgValTrrpProGlnAl 485
1204 TCTACAGCAGGTGACCCCTTATGCTCTTGGTTCGAGTTTGGCCAAAGC 1253
485 aIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysA 502
1254 CATACCTCAGTTCCTGCTGAGACATCTGATCTTCTGGAAGCCGCAAAAG 1303
502 IaLaMetAsnAspAsnGlyLeuGlyLeuPheLeuGlyLysAsnTyr 518
1304 CTGCTCGTGAAGCGAGGTCTACAGGCGCTCTTCTTAGAGGAACTAT 1353
519 ValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAl 535
1354 GTTGGCAGGAGTTCCTGGGCAATGCGTTGGAGGCGGCTATGAAAGATGC 1403
535 aSerGlyValIThrGlyPheLeuSerArgTyrAlaTyrLys 548
1404 CTCGCAAAATATCTGACTTCTTGACCAAGTATGCTCAACAG 1443

seq_name: gb_pat1:A78959
seq_documentation_block: 1698 bp DNA PART 20-OCT-1999
LOCUS A78959
DEFINITION Sequence 13 from Patent M09829554.
ACCESSION A78959
VERSION A78959.1 GI:6092087
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

REFERENCE Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
1 (bases 1 to 1698)
AUTHORS Ishige, F. and Sato, R.
TITLE METHODS OF CONFERRING PRO-INHIBITING HERBICIDE RESISTANCE TO PLANTS
BY GENE MANIPULATION
JOURNAL Patent: WO 9829554-A 13 09-JUL-1998;
ISHIGE FUMIHARU (JP); SATO RYO (JP)
FEATURES
source
1. 1698
/organism="Zea mays"
/strain="B73 INBRED"
/db_xref="taxon:4577"
<2. 1453
/codon_start=1
/product="PROTOPORPHYRYNOGEN OXIDASE"
/protein_id="CA59133.1"
/db_xref="GI:6092088"
/translaton="NSADCVYVGGISGLCTAQAALATRRGVDPVLTENARAPGGNTT
TVRRPERGYLMEBGPNSFQSPDVLITMAVDSGLKDDLVGDPNAPFVLEMEKLRV
SRPADLPEFPLMSIPKLRAGALGIRPPEBRESVEPVRNRTGAEVPEFRLEPE
CSGVYAGDPEKLSMKAAFGWVRLFEETGSIIGTCTKTTOEFSKNKPPRDRARL
GRTVAASRRKLAALPNAITSSUSGSKYKLSMKLSTKSDKGVLEETPEBVSVA
KSYIMTPTSVASNIILRPLSSDADLRSFYTPVAAVTSTPKAIRRECLIDELQ
FGQLHPRSGVETLGTIYSSSLFPPRARGRVLYLNTGATNGLVYSKTESLIEVA
VDRDLRKMILNSTAVDPLVLYGVVMPQALIPQELVYGHLDLLEAKAALDRGVDLFLG
GNVAVGALDRCEVAGVESASQISDFLTYAVK"

BASE COUNT 416 a 396 c 454 g 432 t
ORIGIN

alignment_scores:
Quality: 1978.00 Length: 480
Ratio: 4.455 Gaps: 2
Percent Similarity: 92.500 Percent Identity: 78.750

alignment_block:
US-09-508-418-2 x A78959 ..
Align seg 1/1 to: A78959 from: 1 to: 1698

73 AspCysValIleValGlyAlaGlyIleSerGlyLeuCysIleAlaGlnVa 89
|||||.....
11 GACTGCGTCTGTGGTGGGAGCATTCACTGCGCTCGACACCGCGCGGC 60
89 ImetSerAlaAsnTyr.....ProAsnLeuMetValThrGlnAlaArgA 104
61 GCTGGCCACCGCGGCGACGCGTCTGTCACGAGCGCCCGC 110
104 sPArGAlaGlyLysAsnIleThrThrValGluArg.....AspGlyTyr 118
111 CCGCGCGCGCGGCAACATATACACCGCTGAGCGCCCGCGAGGAAGGGTAC 160
119 LeuTrrpGluGluGlyProAsnSerPheGlnProSerAspProMetLeuTh 135
161 CTCTGGGAGGAGGGTCCCAACAGCTTCCAGCCCTCCAGCCCGCTTCTCAC 210
135 ImetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAspProA 152
211 CATGGCCGTGACAGCGGACTGAAGATGACTTGTGTTTGGGACCCAA 260
152 snAlaProArgPheValleuTrrpLysGlyLysLeuArgProValProSer 168
261 ACGCGCGCGCTTCTGCTGAGGAGGGAAGGCTGAGGCGCGCTGCATTC 310
169 LysLeuThrAspLeuProPheAspLeuMetSerIleLeuProGlyLysLe 185
311 AAGCCCGCGACCTCCGCTTCTGATCTCATGAGCATCCACGAGGAGGACT 360
185 uArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGlyHisG 202
361 CAGGCGCGGCTAGCGCGCGTGGCATCCGCGCCCTCCTCCAGAGCGCGCG 410

202 IuGIuSerValGIuGlnPheValArgArgAsnIeuGIyGIuValPhe 218
 |||||.....
 411 AAGAGTCAGTGGAGAGTCTGCGCCGCAACCTCGTGTGAGGCTTT 460
 |||||.....
 219 GluArgIeuIleGIuPheCysSerGIyValTyrValGIyAspProse 235
 |||||.....
 461 GAGCGCTCATGTAGCTTTCTGCTCAGGTGTCTATNCCTGGTATCTTC 510
 |||||.....
 235 TlyLeuSerMetLysAlaAlaPheGIyLysValTyrPlyLeuGIuGIuT 252
 |||||.....
 511 TAACTCAGCATGAAAGCTCATTTGGGAAAGCTTGGCGGTGAAAGAAA 560
 |||||.....
 252 hrGIyGIySerIleIleGIyGIyThrPheLysAlaIleLysGIuArgSer 268
 |||||.....
 561 CTGGAGGTAGATTATTTGGTGAACCATCAAGCAATTCAGAGAGGAGC 610
 |||||.....
 269 SerThrProLysAlaProArgAspProArgLeuProLysProLysGIyGI 285
 :|||:|.....
 611 AAGAAATCCAAACCCAGGAGGATCCCGCTCCGAAAGCCAAAGGGCA 660
 |||||.....
 285 nThValGIySerPheArgLysGIyLeuArgMetLeuProAspAlaIleS 302
 |||||.....
 661 GACAGTTCATCTTCCAGAAAGGCTTGGCCATGCTTCCAAATGCCATTA 710
 |||||.....
 302 erAlaArgLeuGIySerLysIeuLysIeuSerTyrPlyLysIeuSerIle 318
 :|||:|.....
 711 CATCCAGCTTGGTGAAGTCAACTATCAATGGAACCTCACGAGCATTT 760
 |||||.....
 319 ThrLysSerGIuLysGIyTyrHisLeuThrTyrGIuThrProGIuGI 335
 |||||.....
 761 ACAAAATCAGATGACAAAGGATATGTTGGAGTAAAGCCCAAGG 810
 |||||.....
 335 yValValSerIeuGIuSerArgSerIleValMetThrValProSerTyrV 352
 |||||.....
 811 GGTGTCTTGGTGGTAAAGCTTATCACTATGATCCATTCATCATATG 860
 |||||.....
 352 aAlaSerAsnIleIeuArgProLeuSerValAlaAlaAlaAspAlaLeu 368
 |||||.....
 861 TTGCTAGCAACATTTTGGCTCCACTTTCAAAGCCGATGCTGGAGATGCT 910
 |||||.....
 369 SerAsnPheTyrTyrProProValGIyAlaValHisIleSerTyrProGI 385
 :|||:|.....
 911 TCAAGATTCTATTATCCACCGTGTGCTGTAACTGTTCTGATTCCAA 960
 |||||.....
 385 nGluAlaIleArgAspGIuArgLeuValAspGIuLeuLysGIyPheG 402
 :|||:|.....
 961 GGAAAGCAATTAAGAAAAGATGCTTAATGATGGGAACTCCAGGGCTTTG 1010
 |||||.....
 402 IyGlnIeuHisProArgThrGlnGIyValGIuThrIleGIyTyrIleTyr 418
 |||||.....
 1011 GCCAGTTGCATCCACGTAGTCCAGGAGTTGAGACATTAGAAACAATATAC 1060
 |||||.....
 419 SerSerSerIeuPheProAsnArgAlaProLysGIyArgValIleuLeu 435
 |||||.....
 1061 AGTTCCTCACCCTTCCAAATCGTCTCCAGCGTACGGTGTACTTCT 1110
 |||||.....
 435 uAsnTyrIleGIyGIyAlaLysAsnProGIuIleIeuSerLysThrGIus 452
 |||||.....
 1111 AAACCTATAGGAGGTCTACAAACACAGCAATTTGTTCCAAAGACTGAAA 1160
 |||||.....
 452 erGlnIeuValGIuValValAspArgAspIleuArgLysMetLeuIleLys 468
 :|||:|.....
 1161 GTGAGCTGTGGAGACAGTGAACCGTGAACCTCCGAAAAATGCTTATTAAT 1210
 |||||.....
 469 ProLysAlaGlnAspProLeuValValGIyValArgValTyrProGlnAl 485
 |||||.....
 1211 TCTACAGCAGTGGACCTTTAGTCCCTTGGTGTGAGGTTGGCCACAAGC 1260
 |||||.....
 485 aIleProGlnPheLeuValGIyHisIleuAspThrIleuSerThrAlaLysA 502
 |||||.....
 1261 CATACCTCAGTTCCTGTAAGACATCTTGAATCTTGGAAAGCCGCAAAAG 1310
 |||||.....
 * 502 IaAlaMetAsnAspAsnGIyLeuGIuGIyLeuPheLeuGIyGIyAsnTyr 518

|||||.....
 1311 CTGCCCTGGACCGAGGTGGCTACAGATGGGCTGTCTTCCATGAGGAACTAT 1360
 |||||.....
 519 ValSerGIyValAlaLeuGIyArgCysValGIuGIyAlaTyrGIuValAl 535
 |||||.....
 1361 GTTGGAGGAGTGGCCCTGGGCGAGATCGTGTGAGGGCCGCTATGAAAGTGC 1410
 |||||.....
 535 aSerGIuValThrGIyPheLeuSerArgTyrAlaTyrLys 548
 |||||.....
 1411 CTGGCAAAATATCTGACTTCTTGAACCAAGTATAGCTTACAAAG 1450

OM of: US-09-508-418-2 to: N_Geneseq_0601.* out_format : pfs

Date: Jul 3, 2001 11:40 AM

About: Results were produced by the GenCore software, version 4.5, Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

MODEL=Iramet_p2n.model -DEV=xlp
-Q/cgr2_1/USPTO_spol/US09508418/r/unat_03072001_100017_24474/app_query.fasta_1.613
DB=N_Geneseq_0601 -QPMF=fastap -SUFFIX=ring -GPOP=12.000
GABEPT=4.000 -MINMATCH=0.100 -IOOPCI=0.000 -IOOPEXT=0.000
QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
FGAPOP=6.000 -DELPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-TRANS=human40.cgi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFPM=pfis
-NOM=ext -MILEN=0 -MAXLEN=200000000
-USR=US09508418_eCGN1_1_222 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLIPX
-WALT -THRAD=1

Search information block:

Query: US-09-508-418-2
Query length: 548
Database: N_Geneseq_0601.*
Database sequences: 730101
Database length: 313950809
Search time (sec): 112.980000

score_list:

Table with columns: Sequence, Strd Orig, ZScore, EScore, Len, Documentation, and alignment scores. It lists various sequences and their corresponding scores and lengths.

Table with columns: seq_name, seq_documentation_block, ID, and sequence. It lists sequences like /SIDS8/gcgdata/geneseq/NA1999.DAT and their IDs.

seq_documentation_block:
cdna to mRNA: 1874 BP.
AAAX28279:
16-JUN-1999 (first entry)
N. tabacum protoporphyrinogen oxidase mutant coding sequence.
Protoporphyrinogen oxidase; light-requiring herbicide tolerance; enzyme; ds.
Nicotiana tabacum.
MO9913087-AI.
18-MAR-1999.
10-SEP-1998; 98WC-JP04064.
11-SEP-1997; 97JP-0265084.
(NIPPY) NIPPON NOYAKU CO LTD.
Hirooka T, Horikoshi M, Mametsuka K;

Example 5; Page 44-49; 56pp; Japanese.
This sequence encodes the Nicotiana tabacum protoporphyrinogen oxidase of the invention, which is tolerant to light-requiring herbicides. The enzyme and its derivatives can be applied to construct plants with high tolerance to light-requiring herbicides, useful in agriculture.
Sequence 1874 BP; 509 A; 382 C; 461 G; 522 T; 0 other;

alignment_scores:
Quality: 2836.00 Length: 548
Ratio: 5.185 Gaps: 0
Percent Similarity: 99.818 Percent Identity: 99.818

alignment_block:
US-09-508-418-2 x AAAX28279
Align seg 1/1 to: AAAX28279 from: 1 to: 1874
1 MethrhrtrhrProilealaasnhisProasnllephethrhinsglnse 17
|||||
26 ATGCAACAACCTCCCAATGCCAATCAACCCATTAATTTTCACTGACGCGTC 75
17 rSerSerSerProleualaphelbeuasnAlygThSerPheIleProPhes 34
|||||
76 GTCGCATCCGCAATGGCATTTAAACCCGTAACGATTCATCCCTTTCT 125
34 erSerIleSerIlyvArGAnsSerValAsncysAsnGLYTPArGThrArg 50
|||||
126 CTTCAATTCACAAAGCGCAATATGTGTCAATTCGAATGGCTGGAGAAKACGA 175
51 CysSerValAlaIlyAspPlyrThrValProSerSerAlaValAspIlygl 67

```

|||||
176 TGCTCCGTTGCCAAAAGATTACAGTTCCCTTCACAGGGGCTCGACGGGG 225
67 YPRoAlaIaGIuLeuAspCyValIleValIGlyAlaGIyIleSerGIYL 84
|||||
226 ACCGCCCGGAGCTGGACGTGTATATAGTGGAGCAGGAATTTGGGCC 275
84 eUCyIIeAlaGIInValMetSerAlaAsnTYrProAsnLeuMetValThr 100
|||||
276 TCTGGATTGGCCAGGTTGATGTCCCGTAAATTTACCCCAATTTGATGTAA 325
101 GIuAlaArGAspArGAlaGIyGIyAsnIleThrThrValGIuArGAspGI 117
|||||
326 GAGGGAGAGATCCGCGGTGGCAACATAACGACTGTGGAAGAGACGG 375
117 YTYrLeuTYrPGLuGIuGIyProAsnSerPheGIInProSerSPrPheM 134
|||||
376 CTAAATTTGGGAAAGGTTCCCAACAGTTTCCACCCGCTCGATCCTATGT 425
134 eUThrMetAlaValAspCySGLyLeuLYsAspAspLeuValLeuGIYAsp 150
|||||
426 TGAGTATGGCAGTAGATTTGTGATGAAGATTTTGTGTGGAGAT 475
151 PRoAsnAlaPrOArGpHeValLeuTYrLySGLyLYsLeuArGPrOValPr 167
|||||
476 CCTAATGGCCCGCCGTTTCGTTTGGGAAGGTAATTAAGCCCGCTCC 525
167 oSerLYsLeuThrAspLeuPrOpePheAspLeuMetSerIleProGIYL 184
|||||
526 CTCAAAACCTCACTACTATCTTCCCTTTTGGATTTGATGAGCATTCCTGG 575
184 ySLeuArGAlaGIyPheGIYProIleGIyLeuArGPrOserProProGIY 200
|||||
576 AGTTGAGAGCTGTGGTTTGTGTCATTTGGCCCTCCGCTTCACCCAGG 625
201 HtSGluGIuSerValGIuGIInPheValArGArGAsnLeuGIyGIuVA 217
|||||
626 CATGGAGAACTAGTGGAGCAGTTCGTCGTGTAATCTTGGTGGCAGAG 675
217 IPhelGIuArGIleuIleGIuPrOpePheCySerGIYValTYrValGIYAsp 234
|||||
676 CTTTGAACCGCTTGTAGAACCACTTTTGTGTTTATGTGGTATGTC 725
234 roSerLYsLeuSerMetLYsAlaIaPheGIYValTYrLYsLeuGIu 250
|||||
726 CCTGAAAACCTGATGATGAAGCAGATTTGGAAAGTTTGGAAAGTTGGA 775
251 GIuThrGIyGIYSerIleIleGIyGIYThrPheLYsAlaIleLYsGIuAr 267
|||||
776 GAAACTGGTGTAGCATTTATTTGGAGAACCTTTAAAGCAATTAAGAGAG 825
267 gSerSerThrProLYsAlaPrOArGAspPrOArGLeuProLYsProLYSg 284
|||||
826 ATCCAGTACACCTTAAAGCCCGCGGATCCGCGTTTACCTTAAACCAAAAG 875
284 IYGIInThrValGIYSerPheArGIySGLyLeuArGMetLeuProAspAla 300
|||||
876 GACAGACAGTGGAMTCATTCAGGAAGGCTCAGAAATGCTCCGGAGTCA 925
301 ILeSerAlaArGLeuGIYSerLYsLeuLYsLeuSerTYrLYsLeuSerSe 317
|||||
926 ATCAGTGCMAAGATTGGCAAACTTAAACTATCATGCAAGCTTCTAG 975
317 IILeThrLYsSerGIuLYsGIYLYrHtSLeuThrTYrGIuThrPr 334
|||||
976 CATTACTAAGTCAAGAAAAGAGATATCACTGTACATACAGACACACAG 1025
334 IuGIYValIValSerLeuGIInSerArGSerIleValIleMetThrValProSer 350
|||||
1026 AAGGAGTAGTTCCTTCAAAAGTGAAGCAATTTGATGACTGTCCCAATCC 1075
351 TYrValAlaSerAsnIleLeuArGPrOleuSerValIaIaAlaAspAl 367
|||||

```

```

1076 TATGTAGCAAGCAACATTTACGTCTTTCGGTTCGCCGACAGATGC 1125
367 AlLeuSerAsnPheTYrTYrProProValGIYAlaValIThrIleSerTYrP 384
|||||
1126 ACTTCAAAATTTCTACTATTCGCCAGTGTGAGCAGTCAACAATTTCTATATC 1175
384 roGIInGIuAlaIleArGAspGIuArGLeuValAspGIYIleuLeuLYsGIY 400
|||||
1176 CTCAGAGAGTATTCGTGATGAGCCGTCTGTTGATGTGAACATAAAGGGA 1225
401 PheGIYGIInLeuHtSPrOArGThrGIInGIYValGIuThrLeuGIYThrTI 417
|||||
1226 TTTGGCCAGTTTGCATCCAGTACACAGGAGGTGGAACCACTGTGAAAGAT 1275
417 eTYrSerSerSerLeuPheProAsnArGAlaPrOlySGLYArGValLeuL 434
|||||
1276 ATATAGTTCATTCACCTCTCCCTAACCCGTGCCCCAAAAGGTGGGAGCTAC 1325
434 eUleuAsnTYrIleGIYAlaIaLYsAsnProGIInIleLeuSerLYsThr 450
|||||
1326 TCTTGAACCTAATTTGGAGGAGCAAAAATCCGAAATTTTGTCTAAGACG 1375
451 GIuSerGIInLeuValGIuValValAspArGAspLeuArGIYMetLeuII 467
|||||
1376 GAGAGCCAACTTGTGAGAGTAGTGTGATCGTACCTCAGAAAATAGCTTAT 1425
467 eLYsPrOlySAlaGIInAspPrOleuValIaGIYValArGValTYrPrOG 484
|||||
1426 AAAACCCAAAAGCTCAAGATCCCTTGTGTGGTGTGGCAGATRTAGCCAC 1475
484 IAlaIlePrOGInPheLeuValGIYHtSLeuAspThrLeuSerThrala 500
|||||
1476 AMGCTATCCACAGTTTTTGGTTGTGTCATCTGGATACGCTAAGTACTGCA 1525
501 LYsAlaIaIleMetAsnAspAsnGIYLeuGIuGIYLeuPheLeuGIYLYs 517
|||||
1526 AAAGCTGTATGATGATGATATATGAGGCTTGAAGGCTGTTCCTTGGGGTAA 1575
517 nTYrValSerGIYValIaIleGIYArGcySValGIuGIYAlaTYrGIuV 534
|||||
1576 TTTATGTGTCAAGGTGTAGCATTTGGGAGGTTGTGTGTAAGTCTTATGAAG 1625
534 AlAlaSerGIuValIThrGIYpHeLeuSerArGIYrAlaTYrLYs 548
|||||
1626 TTGCATCCGAGTAAACAGGATTTCTGTCTCGGATATCAATACAAA 1669
seq_name: /SIDS8/gcgdata/geneseq/geneseq/NA1999.DAT:AAx28278
seq_documentation_block:
ID AAx28278 standard: cDNA to mRNA; 1874 BP.
XX AAx28278;
DT 16-JUN-1999 (first entry)
DE N. tabacum protoporphyrinogen oxidase coding sequence.
KW Protoporphyrinogen oxidase; light-regulating herbicide tolerance; enzyme;
KW ds.
OS Nicotiana tabacum.
PN W09913087-A1.
PD 18-MAR-1999.
XX 18-MAR-1999.
XX 10-SEP-1998; 98WO-JP04064.
XX 11-SEP-1997; 97JP-0265084.
PR (NIPPON NOYAKU CO LTD.
XX (NIPPON NOYAKU CO LTD.
PA Hirooka T, Horikoshi M, Mametsuka K;
XX PI

```

XX MPI: 1999-215067/18.
 DR P-PSDB: AAY05200.
 XX
 PT Protoporphyrinogen oxidase tolerant to light-requiring herbicides,
 XX useful for constructing plants with such properties
 XX
 XX Example 5; Page 34-39; 56pp; Japanese.
 CC This sequence encodes the Nicotiana tabacum protoporphyrinogen oxidase
 CC of the invention, which is tolerant to light-requiring herbicides. The
 CC enzyme and its derivatives can be applied to construct plants with high
 CC tolerance to light-requiring herbicides, useful in agriculture.
 XX
 XX Sequence 1874 BP: 509 A; 383 C; 461 G; 521 T; 0 other:

alignment_scores:
 Quality: 2832.00 Length: 548
 Ratio: 5.177 Gaps: 0
 Percent Similarity: 99.818 Percent Identity: 99.635

alignment_block:
 US-09-508-418-2 x AAY28278 ..

Align seg 1/1 to: AAX28278 from: 1 to: 1874

1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnse 17
 |||||||
 26 AIGACACACACCTCCATCCGCAATCCCTAAATATTTCTACGCACAGTC 75
 |||||||
 17 rSerSerSerProLeuAlaPheLeuAsnArgrThrSerPheIleProPhes 34
 |||||||
 76 GFCGTACFCGCCCATTTGGCATTTAAACCCGTACGAGTTTCATCCCTTCT 125
 |||||||
 34 eIerIleSerIysArgAsnSerValAsnCysAsnGlyTrpParGrHrArg 50
 |||||||
 126 CTTCAACTCCCAAGCGCAATATGTCATATTCATATTCGCAATGGCGGAGAACCGA 175
 |||||||
 51 CysSerValAlaIlysAspTrpThrValProSerSerAlaValAspIlyGI 67
 |||||||
 176 TGCTCCGTTGCCAAAGATTTACACAGTTCCCTCCAGCGGTGCACGGCGG 225
 |||||||
 67 yProAlaAlaGluLeuAspCysValIleValAlaGlyAlaGlyIleSerGIYL 84
 |||||||
 226 ACCCGCGCGGACGCTGCTGTATTAATGGTAGCAGGAATTAATGAGCC 275
 |||||||
 84 eucystrIleAlaGlnValMetSerAlaAsnTrpProAsnLeuMetValThr 100
 |||||||
 276 TCTGCATTTGGCGAGGTGATGTCGCTAATTAATCCCAATTTGGATGTAACC 325
 |||||||
 101 GluAlaArGAspArGAlaGlyIlyAsnIleThrThrValGluArGAspGI 117
 |||||||
 326 GAGCGGAGAAATGCTGCGCGGTGGCAACATACACAGCTGTGAAAAGAGACGG 375
 |||||||
 117 yTrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134
 |||||||
 376 CTATTTTGGGAAGAAGTCCCAACAGTTTCCAGCCGTCGATCTATAT 425
 |||||||
 134 eurThrMetAlaValAspCysGlyLeuIlyAspAspLeuValIleuGIyAsp 150
 |||||||
 426 TGAGCTATGGAGAGATGGATGGATGGAAAGATGATGGTGGGAGAT 475
 |||||||
 151 ProksnAlaProArGrPheValIleuTrpIlyGlyIlyLeuArGrProValPr 167
 |||||||
 476 CCFRAATGCCGCCCTTTCTTTTGTGGAAAGGTAATAAATTAAGGCCCGCTCC 525
 |||||||
 167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGIYL 184
 |||||||
 526 CTCAAAACCTCACTGATCTCCCTTTTGGATTTGATGGAGCATTCCTGGCA 575
 |||||||
 184 yIleuArGAlaGlyPheGlyProIleGlyLeuArGrProSerProProGIYL 200
 |||||||

576 AGTTGAGAGCTGGTTTTGGTCCCATTTGGCCCTCCCGCCCTTCCACCTCCAGGT 625
 |||||||
 201 HSIgluIleuSerValIgluIlePheValArGrArGAsnLeuGIyGIyIyVa 217
 |||||||
 626 CATGAGAAATCAGTTTGACAGATGTCGCTGCTAATCTGGTGGCGGAAGT 675
 |||||||
 217 lPheGluArGLeuIleGluProPheCysSerGIyValTrValGIyAspP 234
 |||||||
 676 CTTTGAAACGCTTGATAGAACCATTTTGTCTGGTCTTAACTCTGATGATC 725
 |||||||
 234 roSerLysLeuSerMetIlySAlaAlaPheGIyIyValITrPlyIleuGIu 250
 |||||||
 726 CCTCAAAACCTGATGTAAGCAGCATTTGGGAAAATTTGGAAAGCTTGGAA 775
 |||||||
 251 GluThrGIyGIySerIleIleGIyIyThrPheIySAlaIleIyGIyAr 267
 |||||||
 776 GAAACTGGTGGTACATTTATTTGGAGAACCTTTAAAGCAATAAAGGAGAG 825
 |||||||
 267 gSerSerThrProIySAlaProArGAspProArGrLeuProIySProIySG 284
 |||||||
 826 ATCCAGTACACCTAAAGCGCCCGCGGATCCGCGTTTACCTFAAACCAAAAG 875
 |||||||
 284 IyGIuIThrValGIySerPheArGlyIyLeuArGMetLeuProAspAla 300
 |||||||
 876 GACAGACAGTTGGATCATTTCAGAAAGGCTCAGAAATCTGCTCCCGATGCA 925
 |||||||
 301 lIeSerAlaArGLeuGIySerIySLeuIySLeuSerTrPlyIyLeuSerSe 317
 |||||||
 926 ATCACTGCAGAGTTGGAAAGCAAAATTAACATATCATGGAAAGCTTTCTAG 975
 |||||||
 317 rIleThrIySserGIuIyGIyIyTrHisLeuThrTrGIuIThrProg 334
 |||||||
 976 CATTACTAAGTCAAAAAAGGAGATATCACTTGCATCAATACAGACACACAG 1025
 |||||||
 334 IuGIyValIleuSerIleuGInserArGserIleValMetThrValProSer 350
 |||||||
 1026 AAGGAGTGTCTCTCAAGTGCAGAGCATTTGTATGATGTCATGTCATCC 1075
 |||||||
 351 TrValAlaSerAsnIleleuArGrProLeuSerValAlaAlaAlaAspAl 367
 |||||||
 1076 TATGTAGCAGACACATATTAAGTCTCTTTTCGGTGGCGGAGCAGATGC 1125
 |||||||
 367 aLeuSerAsnPheTrIyTrProProValAlaValAlaThrIleSerTrP 384
 |||||||
 1126 ACTTTCAAAATTTCTATATCCCGCCAGTTGGAGCAGTCAAAATTTCAATTC 1175
 |||||||
 384 roGInGluAlaIleArGAspGIuArGLeuValAspGIyGIuIleuIyGIy 400
 |||||||
 1176 CFCAGAGACGTATGCTGATGAGCGGTCTGTTGATGATGGAACTTAAGGGA 1225
 |||||||
 401 PheGIyGlnIleuHisProArGrThrGInGIyValGIuIThrLeuGIyThrI 417
 |||||||
 1226 TTTGGGAGTTGCATCCACGTACACAGGAGGAGTGAACACATPAGAACGAT 1275
 |||||||
 417 eTrSerSerSerLeuPheProAsnArGAlaArGIySGLyArGIyValIleu 434
 |||||||
 1276 ATATATGTTTCATCACTCCCTTAACCGGTGCCCAAAAAGTGGGGTCTAC 1325
 |||||||
 434 eIleuAsnTrpTrIleGIyIyAlaIlyAsnProGIuIleuSerIySThr 450
 |||||||
 1326 TCTTGAATCTACTTGGAGGAGCAAAAATTCCTGAAATTTTGTCTPAGACG 1375
 |||||||
 451 GluSerGlnIleuValAlaIyValAspArGAspLeuArGIySmetLeuI 467
 |||||||
 1376 GAGAGCAACCTTGTGGAAAGTATGATGCTGCTCAGAAAAAATCTTAT 1425
 |||||||
 467 eIySProIySAlaGlnAspProLeuValIyAlaIyValArGIyValITrProg 484
 |||||||
 1426 AAAACCCAAAGCTCAAGATCTCTTGTGGTGGAGTGGAGTGGCCAC 1475
 |||||||
 484 lAlaIleProGInPheLeuValGIyHisLeuAspTrIleuSerThrAla 500
 |||||||
 1476 AAGCTATCCACAGTTTTTGGTGGTCACTGAGATACCTAAGTACGTAC 1525
 |||||||

```

501 LysAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAs 517
    |||
1526 AAGCTGATGAAATGAAATGATGCGCTTAAAGCGCTTCTGGGGGTA 1575
517 nTyrValSerGlyValAlaLeuGlyArgCysValGluGlyValAlaTyrGluV 534
    |||
1576 TTATGTGCAGGTGTAGCATTTGGGAGGAGGTGTGCAAGCTCTATGAA 1625
534 AlaIleSerGluValThrGlyPheLeuSerArgTyrAlaTyrIlys 548
    |||
1626 TTGCATCCGAGGTAACAGGATTTCTGCTCGGTATGCAATACAAA 1669
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04313
seq_documentation_block:

```

```

ID AAV04313 standard; cDNA; 1826 BP.
XX
AC AAV04313;
XX
DE 20-APR-1998 (first entry)
XX
XX Cotton protox-1 cDNA.
XX
KM Protoporphyrinogen oxidase-1; protox-1; promoter; cotton;
KM herbicide resistance; breeding programme; probe; gene isolation;
KM genomic mapping; ss.
XX
OS Gossypium hirsutum.
XX
FH Key location/Qualifiers
FT CDS 31..1650
FT FT /*tag= a
FT FT /product= protox-1
XX
XX MO9722028-A1.
XX
XX 04-SEP-1997.
XX
XX 27-FEB-1997; 97MO-US03343.
XX
XX 21-JUN-1996; 96US-0020003.
XX 28-FEB-1996; 96US-0012705.
XX 28-FEB-1996; 96US-0013612.
XX
XX (NOVS ) NOVARTIS AG.
XX
XX Johnson MA, Volrath SL, Ward ER;
XX
XX WPI; 1997-489209/45.
XX
XX P-PSDB; AAM41609.
XX
XX DNA containing a plant proto-porphyrinogen oxidase gene promoter -
XX optionally linked to a heterologous gene, especially to express
XX herbicide-resistant enzymes, and plants containing such constructs
XX
XX Disclosure; Pages 78-80; 114pp; English.
XX
XX The present sequence encodes cotton protoporphyrinogen oxidase-1
XX (protox-1).
XX The protox-1 promoter can be used to express herbicide resistant
XX enzymes, specifically protox, i.e. a plant tissue, plant or progeny
XX containing a chimeric gene of the promoter and a heterologous
XX coding sequence. The plant can also be used in breeding programmes.
XX Also hybridising fragments of the protox coding sequence can be
XX used as probes, e.g. to isolate related genes or for genomic
XX mapping.
XX
XX Sequence 1826 BP; 471 A; 399 C; 439 G; 517 T; 0 other;

```

```

alignment_scores:
Quality: 2113.00 Length: 506

```

```

Ratio: 4.477 Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249
alignment_block:
US-09-508-418-2 x AAV04313 ..
Align seg 1/1 to: AAV04313 from: 1 to: 1826

```

```

47 TrpArgThrArgCysSerValAlaIleValAspTyrThrValProSerSerAl 63
    |
130 TTCAAGCTCCGATGCTCCCTCCGCCAGGCTCCACGATTTCCATCTTAA 179
63 aValAspGlyGlyProAlaIleValLeuAspCysValIleValAlaGlyAla 80
    |
180 AATCGAAGCGGGGAGAAATCCATCCCGGATTCGGTATCCGATCGTTGGAG 229
80 IylIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
230 GTATACAGTGCATTTCCATTCACAAAGCTCTCCGCCCAAGCACCCTGAC 279
95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
280 GTCCGCTTCCATATGTGATGTGACGGAGGCCAGACCGCTGTGGTGCA 329
109 nIleThrThrValGluIleArgAspGlyTyrLeuTrpGluGlyProAsn 126
330 CATCACTACCGCTTGAGAGATGATATCTGTGGAGAAAGGCCCCACA 379
126 ePheGlnProSerAspPrometLeuThrMetAlaValAspCysGlyLeu 142
380 GTTTTCAGCCCTCCGATTCCTATTCATCAACCATGCGCTGATGTGATG 429
143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
430 AAGGACGATTTGGTTTATGAGTCCCTAATGACCGCGATTTGTACTATAG 479
159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
480 GAGGGGAAAACCTAAGGCCCTGTGCCCAAGCCAAAGCCGATTCGGCTTT 529
176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
530 TTGATTTGATGAGCATTGCTGMAAACTTAAGGCTGTCCGGGCTATT 579
193 GlyLeuArgProSerProProGlyHisGluGlnSerValGluGlnPheVa 209
580 GCCATTCGGGCTCCCTCCCGGTTATGAAAGATTCGGTGGAGGAGTTTGT 629
209 IArgArgAsnLeuGlyGlyValPheGluArgLeuIleGluProPheC 226
630 GCSCCGTAATCTTGAGGCTGAGGTTTGAACGCTTATTTGAACCATTTT 679
226 ySerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
680 GTTCAGGTGTATTATGCAAGGGGATCTTCAAAATTAAGCAATGAAACGAG 729
243 PheGlyLysValTyrPheLysLeuGluGlnThrGlyLysSerIleIleGly 259
730 TTTCGAAGAGTATGAGAGCTTAGAAGATTTGGTGGCAGCATATTTGGT 779
259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
780 CACTTCAAGACAAATCCAGAGAGAAATTAAGACACCTTAAGCCACCAAG 829
276 sProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
830 ACCCGCTCCCAAAACCGAAGGCCAAACAGTTGAGCTTTTGAAGAG 879
293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
880 GGACTTACCATGCTGCTGAGGCAATTCGTAACAGTTTGGTAGCAAGT 929
309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326

```



```

1180 TTTGATTCAGATGTCGAACTTAAGGGGGTTGGCCAGTTGCACCACCGAGCCCA
409 nGlyValIGluThrLeuGlyThrIleTyrSerSerLeuPheProAsnA
1230 AGGAATTTGAAACTTTAAGGAGATATACAGTTCATCACCCTTCCCAATC
426 rGAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys
1280 GAGCTCCATCTGGCAGGGTTCCTCTTGAACACTAGAGAGAGACTACC
443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValAla
1330 AACCTGGAATTTTGTCCAAAGACTGAAGGGGAACTGTGTGAAGCAGTTGA
459 PATGAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV
1380 TCGTGAATTTGAGAAAATGCTTAATAATCCCTAATGCAAAAGATCCCTGTG
476 aYValGlyValArgValITrProGlnAlaIleProGlnPheLeuValGly
1430 TTTTGGGTGTAAAGATATGGCCAAAAGCCATTCACAGTTCCTGGTGGT
493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe
1480 CATTTGGATCTCCTGTGATGTGCAAAAATGCGCTCAGGGGATTCGGGTT
509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA
1530 TCAMGAGCTTTCTTGGGGCAACTATGTATCTGTGTGCAATTAAGGAC
526 rGcysValIGluGlyAlaTyrGluValAlaSerGluValIThrGlyPheLeu
1580 GGTGTGTGGAGAGGTGCTTAGAGGTTGCAAGCTGAAGTGAAGAAATTCCTG
543 SerArgTyrAlaTyrLys 548
1630 TCACAAATATGCAATACAAA 1647

```

```

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76579
seq_documentation_block:
ID   AAF76579 standard; cDNA; 1826 BP.
XX
XX   AAF76579;
AC
XX
DE   16-MAY-2001 (first entry)
XX
XX   Cotton protoporphyrinogen oxidase coding sequence SEQ ID NO: 15.
XX
XX   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein; ss.
OS   Gossypium hirsutum.
XX
XX   WO200112825-A1.
XX
XX   22-FEB-2001.
XX
XX   30-JUN-2000; 2000WO-EP06127.
XX
XX   13-AUG-1999; 99US-0373691.
XX
XX   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
XX   Johnson MA, Volrath SL, Helfetz PB, Law MD;
XX
XX   WPI: 2001-234914/24.
XX
XX   P-PSDB: AAB72907.
XX
XX   Plant DNA molecules encoding herbicide-tolerant forms of
XX   protoporphyrinogen oxidase which are useful for rationally designing
XX   new inhibitory herbicides and for producing herbicide-tolerant
XX   transgenic plants and seeds -

```


110 ILehrThrValAGluArgAspGlyTyrLeuThrPrpGluGluGlyProAsnSe 126
 325 APTATCACTCGTAGAGAAATGGTTTCTGTGGGAGAAAGGTCCCAATAG 374
 126 PheGlnProSerAspPrometLeuThrMetAlaValAspCysGlyLeuL 143
 375 TTTTCAACCGTCTGATCTATGCTCAGTATGTGTGTAGATGTGGTTTGA 424
 143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
 425 AGGATGATTTGGTGTGGAGATCCTACTGCGCCAAAGGTTTGTGTGG 474
 160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
 475 AATGGGAAATTTGAGCCCGTTCATCGAAGCTAACAGACTTACCGTTC 524
 176 eaSPLeuMetSerIleProGlyLysLeuArgAlaIlePheGlyProIleG 193
 525 TGATTTGATAGATATTGGTGGAGATTAAGAGCTGGTTTGGTGCAC 574
 193 LysLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
 575 GCATTCGACCGTCACCTCCAGGCTGGAAGAAATCTGTGGAGAGATTGTA 624
 210 ArgATGAsnLeuGlyGlyValValPheGluArgLeuIleGluProPheCy 226
 625 CCGCGTAACTCGGTGATGAGAGGTTTTTGAAGCCCTGATTTGAACCG 674
 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaIleP 243
 675 TTCACGCTGTTTCTCGTGTGATCCTTCAAAACTGAGCATGAAAGCGCG 724
 243 heGlyLysValThrPrpLysLeuGluGluGlySerIleIleGlyLys 259
 725 TTGGGAAAGGTTTGGAACTAGAGCAAAAGTGGAGAACATTAATAGTGT 774
 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
 775 ACTTTTAAGGCCAATTCAGGAGAGAAAAGCCCTCCCAAGGCCAAGACGA 824
 276 ProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
 825 CCCGGCCCTGCCAAAACACAGCCCAAAACAGTTGTTCTTCCAGAGAG 874
 293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
 875 GACTTCGATGTTGCCAGAGCAAAATCTGCCAAGATTTAGGTAAGCAAA 924
 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyLys 326
 925 AAGTTGTCTTGGAGAGCTCCAGGTATCACTAAGCTGAGAGCGGAGATA 974
 326 HisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
 975 CAACCTTAACATATGAGACTCCAGATGTTAGTTCCTCGTAGAGCAAAA 1024
 343 eRIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
 1025 GTTGTGTAAATGACGGTCCCATCTCATGTTGCCAAGTGGTCTCTGGCC 1074
 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
 1075 CTTTCTGAATCTGTCGCAAAATGCACTCCAAAACAATATTTCACCCAC 1124
 376 IGLyAlaValThrIleSerTyrProGluGluAlaIleArgAspGluAAGL 393
 1125 TCCACACAGATATCTATCTGTACCCGAAAGAACAAATCCGAACAGATGT 1174
 393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
 1175 TGATAGATGGTGAACATAAGGGTTTGGCCAAATTTGATTCACAGCCGCA 1224
 410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426

1225 GGAGTTGAACATTTAGGAACTATCTACAGCTCCTCACTCTTCCAAATCG 1274
 426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
 1275 CGCACCGCCCGGAAAGAAATTTTGTGTTGAACATTAATGGCGGGCTTACA 1324
 443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValAsp 459
 1325 ACACCGGAATTCGTCCAMGTGAAAGGTGAGTTAGTTCGAGACAGATTGAC 1374
 460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
 1375 AGAGATTTGAGGAAATGCTAATTAAGCCCTAATTCGACCGCATCCACTTA 1424
 476 lValGlyValArgValIleTrpGlnAlaIleProGlnPheLeuValGlyH 493
 1425 ATTAGGAGTTTGGGTAGGCTCAGCCATTCCTCAGTTCATGATGGTGC 1474
 493 lIleuAspThrLeuSerThrAlaLysAlaIleMetAsnAspAsnGlyLeu 509
 1475 ACTTTGATATCTTGCACCGCTAAATATCATCTAATACGCTTCGGGCTAC 1524
 510 GlnGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLysGlyAr 526
 1525 GAAAGGCTATTTTGGGTGGCAATTAACGTCGCTGCTGATAGCCCTTAAGCCG 1574
 526 gCysValAGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
 1575 GTGTGTGAAAGCGCCATATGAACCCGCGATTTGAGGTCAACACTTCAATG 1624
 543 eArgTyrAlaTyrLys 548
 1625 CACGGTACGGCTTACAG 1641
 seq_name: /SIDS8/gcsgdata/geneseq/geneseqn/NA1997.DAT:AA186129
 seq_documentation_block:
 ID AAT86129 standard; cDNA; 1719 BP.
 AC AAT86129;
 DT 01-MAR-1998 (first entry)
 XX
 DE Arabidopsis protoporphyrinogen oxidase (protox-1) cDNA pMDC-2.
 KW Protox-1; protoporphyrinogen oxidase; inhibitor;
 herbicide tolerance; herbicide resistance; transgenic plant; ss.
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT CDS 31..1644
 FT /tag= a
 XX
 PN WO9732011-A1.
 PD 04-SEP-1997.
 XX
 PE 27-FEB-1997; 97WO-US03313.
 XX
 PR 21-JUN-1996; 96US-0020003.
 PR 28-FEB-1996; 96US-0012705.
 PR 28-FEB-1996; 96US-0013612.
 XX
 PA (NOVS) NOVARTIS AG.
 XX
 PI Helfetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER;
 DR MPI:1997-448683/41.
 DR P-PSDB; AAW25746.
 XX
 PT New DNA encoding plant protoporphyrinogen oxidase enzyme - and

herbicide resistant mutants, useful to prepare plants resistant to herbicide which therefore kills undesired vegetation only

Example 2; Page 95-99; 196pp; English. This cDNA clone codes for Arabidopsis protoporphyrinogen oxidase (Protox-1, see AAM25746). The clone has been deposited in pMDC-2 as NRRL B-21238. The isolated cDNA can be mutated so that it encodes a modified protox enzyme that is resistant to protox inhibitors and hence is herbicide tolerant. Plants, especially crop plants, may be engineered for resistance to protox inhibitors via mutation of the native protox gene to a resistant form, or they may be transformed with a gene encoding an inhibitor-resistant form of a plant protox enzyme, including claimed forms from wheat, soybean, cotton, sugarcane, oilseed rape, rice, sorghum and Arabidopsis (see AAM25738-48). Application of herbicide acids can also be used as a vegetation only. Plant protox nucleic acids can also be used as probes and PCR primers, as selectable markers in plant cell transformation methods, and for recombinant production of protox enzymes in host cells. Arabidopsis Protox-1 cDNA has been used in the identification of soybean (see AAT86122), sugarcane (see AAT86124) and rape (see AAT86125) Protox-1 cDNA.

Sequence 1719 BP; 460 A; 351 C; 433 G; 475 T; 0 other; Alignment scores: Quality: 2077.00 Length: 539 Ratio: 4.230 Gaps: 4 Percent Similarity: 91.095 Percent Identity: 75.139

Alignment_block: US-09-508-418-2 x AAT86129 .. Align seg 1/1 to: AAT86129 From: 1 to: 1719

19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
28 FCCARGAGATGATCTCTCCGCGAGACTCAATCGCTTTCGCGC 77
35 rIleSerLysArgAsnSerValAsnGlyTrp.....ArgT 49
78 GTTTTCGAAGCCCAAT..CTCCGATTAATGTTTAAAGCCCTCTTAGC 124
49 hArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
125 FCCGTTGTCAGTGGCGGCTGGACCAACCGCTGGATTCGAAATTCGAA 174
66 GlyGlyProAlaAlaGluLeu.....AspCysValIleValGlyAlaG 80
175 GCGGGGAGGAGCCACCACATACGAGCGATGTGTGATGTGCGGGGAGG 224
80 rIleSerGlyLeuCysIleAlaGluValMetSerIleAsnTyr..... 94
225 TATTAGTGGTCTTTGGCATGCTCAGGCGCTTGTACTAGGACATCCTGATG 274
95ProAsnLeuMetValIleThrGluAlaArgAspArgAlaGlyLys 109
275 CFCGCGCCGAATTTAATTTGTGACCGAGGCTAAGATCGTGTGGAGCAAC 324
110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
325 AATTATACACCTGTAAGAGATGTTTCTCTGGGAAAGAGTCCCAATATG 374
126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 143
375 TTTTCAACCGTGTGATCTTACTCCTCACTAATGTTGTTAGTGGTTTGA 424
143 yAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
425 AGGATGATTTGGTGGTGGAGATCCTACAGCGCAAGTGTGTGGTGG 474
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176

475 AATGGAAATTTAGAGCCGGTTCACATCGAACCTAACACACTTACCGTCTT 193
176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 193
525 TGATTTGATGATGATTTGGTGGAAAGATTTAGACTGCTTTTGGTGCACCTG 574
193 ILeuArgProSerProProGlyHisGluLysLeuSerValGluGlnPheVal 209
575 GCATTCGACCGTCCACCTCCAGGTCGTGAAGATCTGTGGAGGATTTGTA 624
210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
625 CGGGGTAAACCTCCGGTGTGAGGTTTTGGACCGCTGTAAACCGTTTTG 674
226 sSerGlyValIleValGlyAspProSerLysLeuSerMetLysAlaAla 243
675 TTCAGGTGTTTATGCTGTGATCTTCAAAAACGTAGCATTAAGACAGCGCT 724
243 heGlyLysValIleTrpLysLeuGluGluThrGlySerIleIleGlyGly 259
725 TTGGGAAGGTTTGGAAACTAGACCAAAATGTTGGAAAGCATTAATVAGTGGT 774
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg 276
775 ACTTTTAAAGCAATTCAGAGAGGAAAGCAAGCTCCCAAGCAGAGACAGA 824
276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 293
825 CCCGCGCTGCCAAAACACAGGCGCCAAAACGTTGGTCTTTCAGAGAGG 874
293 ILeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
875 GACTTCGAAATGTCCAGAAACCAATATCTCCAAAGATTAAGTACGAAAGTT 924
310 LysLeuSerTrpLysLeuSerSerIleThrLysSerLysGlyGlyTyr 326
925 AAGTTGCTGGAAAGCTCAGGATATCAGTAAAGCTGAGGAGGAGGATA 974
326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 343
975 CAACTTAAACATATGAGACTCCAGATGTTAGTTCCTCCGTCAGACGAA 1024
343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
1025 GTGTTGTAATATGACGGTCCATCTCATGTGTGCAAGTGTCTTCGCGCCT 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTrpProPro 376
1075 CTTTCTGAAATCTGCTGCAAAATGACCTCTCAAAACTATATTTACCCAC 1124
376 IGLValAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 393
1125 TGCGACGATATCTATCTCGRACCCGAAAGAAAGCAATCCGAAACAGATGTT 1174
393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgTyrGln 409
1175 TGATTAAGATGGTAACCTAAGGGTTTTGGCAATTTGCATCCACGACGCA 1224
410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsn 426
1225 GGAATTTGAAACATTAAGAACTATACAGCTCCCTCATCTTTCCAAATTC 1274
426 gAlaProLysGlyValValLeuLeuLeuAsnTyrIleGlyValAlaLys 443
1275 CGCACCCCGCGAAGAAATTTGCTGTGAACTACATTGGCGGCTCTACAA 1324
443 snProGluIleLeuSerLysThrGlnSerGlnLeuValGluValValAsp 459
1325 ACACCGCAATTTCTCCAAAGTCAAGGATTAAGTGAAGGAAAGGTTTGC 1374
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeu 476

1375 AGAGATTGGAGGAAAATGCTAATTAATGACCTTAATTCGACCGATCCACTTAA 1424
 476 IVALGIYVALARGVALTRPPROGLAALAIlePROGLInPheLeuValGLVH 493
 1425 AATTAGAGATTAGGGTATGGCCCAAGCCATTCTCAGTTTCTAGTTGGTCC 1474
 493 ILeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGLYLeu 509
 1475 ACTTTGATATTCCTTGACACGGCTTAATATCATCTCTAAGCTTCGGGGCTAC 1524
 510 GIUGIYLeuPheLeuGLYGLYAsnTYrVALSerGIYVALAlaLeuGLYAr 526
 1525 GAAGGGCTAATTTTTGGTGGCAATTAACCTCGCTGTGTAGCCCTTAGGGCCG 1574
 526 GCYSVALGIUGIYVALATYrGLUVALAlaSerGLUVALThrGIYpHeLeuS 543
 1575 GTGTGTAGAGAGCCGATATGAAACCCGCAATTTGAGGTCAACACATTCATGT 1624
 543 eRArgTYrAlaTYrLys 548
 1625 CACGGTACGCTTACAAG 1641

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV07251
 seq_documentation_block:
 ID AAV07251 standard; cDNA; 1719 BP.
 AC AAV07251;
 XX
 DT 14-AUG-1998 (first entry)
 DE Arabidopsis cDNA encoding proto-porphyrinogen oxidase-1 (protox-1).
 KW Proto-porphyrinogen oxidase; protox; herbicide tolerant; resistance;
 KM Inhibitor; variegate porphyrin; ss.
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT CDS 31..1664
 FT /tag= a
 FT /product= Arabidopsis_protox-1
 XX
 PD US5767373-A.
 XX 16-JUN-1998.
 XX 16-JUN-1994; 94US-0261198.
 XX 06-JUN-1995; 95US-0472028.
 XX 16-JUN-1994; 94US-0261198.
 XX (NOVS) NOVARTIS FINANCE CORP.
 XX
 XX Volrath S, Ward ER;
 DR WPI: 1998-361821/31.
 DR P-PSDB; AAM51347.
 XX
 PT DNA encoding eukaryote herbicide resistant proto-porphyrinogen
 PT oxidase - useful for producing recombinant plants having functional
 PT enzyme, to be grown in the presence of herbicides
 XX
 PS Claim 6; Columns 41-46; 43pp; English.
 XX
 CC The invention relates to eukaryotic DNA sequences coding for native
 CC proto-porphyrinogen oxidase (protox) or modified forms of the enzyme
 CC which are herbicide tolerant. Plants having altered protox activity
 CC which confers tolerance to herbicides are also provided. These plants
 CC may be bred or engineered for resistance to protox inhibitors via
 CC mutation of the native protox gene to a resistant form or through
 CC increased levels of expression of the native protox gene, or they may be
 CC transformed with modified eukaryotic or prokaryotic protox coding

CC sequences or wild type prokaryotic protox sequences which are herbicide
 CC tolerant. In the human condition variegate porphyria, an autosomal
 CC dominant disorder characterised by neuropsychiatric symptoms, decreased
 CC protox activity is detected. Thus the DNA molecules can be used to
 CC provide probes to detect and quantify protox levels in the diagnosis of
 CC diseases associated with the enzyme. The present sequence represents
 CC Arabidopsis cDNA encoding proto-porphyrinogen oxidase-1 (protox-1).
 XX
 SQ Sequence 1719 BP; 460 A; 351 C; 433 G; 475 T; 0 other:
 alignment_scores: Length: 539
 Quality: 2077.00 Gaps: 4
 Ratio: 4.230
 Percent Similarity: 91.095 Percent Identity: 75.139
 alignment_block:
 US-09-508-418-2 x AAV07251 ..
 Align seg 1/1 to: AAV07251 from: 1 to: 1719
 19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
 28 TCCATGGAGTTATCTCTTTCGCGACGACCAATCCGCTTCTCCGTC 77
 35 rIleSerLysArgAsnSerValAsnCysAsnGLYTrp.....ArgT 49
 78 GTTTTCGAAGCCCAAT...CTCCGATTAATYGTTTTAAAGCCCTTAGAC 124
 49 hRArgCYSerValAlaLysAspTYrThrValProSerSerAlaValAsp 65
 125 TCCGTGTTCAGTGGCCGCGTGACCAACCGTCGATCTCAAAATCGAA 174
 66 GLYGLYProAlaAlaGLULeu....AspCysValIleAlaGLYAlaGL 80
 175 GCGGAGGAGGACACACACACACGACGATGTGTGATGTGCGGAGG 224
 80 YIleSerGIYLeuCysIleAlaGLInValMetSerAlaAsnTYr..... 94
 225 TATTAGTGGTCTTTCATCGCTCAGCGCGCTTGGTACTAAGCAATCCGANG 274
 95ProAsnLeuMetValThrGLUAlaArgAspArgAlaGLYIAsn 109
 275 CTGCTCCGAATTAATTTGACCGAGGCTTAAGGATGCTGTGGAGGCAAC 324
 110 IleThrThrValAlaGLUArgAspGLYTYrLeuTrpGLUGLYProAsnSe 126
 325 AATATTCACCTCGTGAAGAAATGTTTCTCTGGGAAAGAGTCCCAATAG 374
 126 rPheGLNProSerAspProMetLeuThrMetAlaValAspCysGLYLeuL 143
 375 TTTTCAACCGCTCGATCCCTATGCTCACTAAGTGTGATAGATAGGTTTGA 424
 143 ysAspAspLeuValLeuGLYAspProAsnAlaProArgPheValLeuTrp 159
 425 AGGATGATTTGGTGGAGATCTTACCTGCCCAAGAGTGTGGTGG 474
 160 LysGLYLYsLeuArgProValProSerLysLeuThrAspLeuProPheh 176
 475 AATGGGAATTTGAGCCCGGTTCCATCGAAGCTAAGACTTACCGTCTT 524
 176 eAspLeuMetSerIleProGLYLYsLeuArgAlaGLYpHeGLYProIleG 193
 525 TGATTTGATGATGATTTGGTGGAAAGATTTAGAGCTGTTTGGTCCACTTG 574
 193 LysLeuArgProSerProProGLYHISGLUgluSerValAluGLInPheVal 209
 575 GCATTTGACCCCTCACCTCCAGGTCTGAAAGAAATCTGTGAGAGATTTTGA 624
 210 ArgArgAsnLeuGLYGLYValAlaPheGLUArgLeuIleGLYLeuProPhe 226
 625 CGCGTAACTCCGATGATGAGGTTTTGGAGCGCTGATTAAGCAACCGTTTGG 674

226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaIleP 243
 675 TTGAGGTGTTTATGCTGGTGTATCTTCAAACTGGACGATGAAAGCGCT 724
 243 heGlyLysValTyrLysLeuGluGluThrGlyGlySerIleIleGly 259
 725 TTGGAAAGGTTGGAAACTAGAGCAAAATGATGGAAAGCATTAATAGTGT 774
 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
 775 ACTTTAAAGCAATTCAGAGGAGGAAAAAGCTCCCAAGGCAACGAGA 824
 276 pProArgLeuProLysGlyGluThrValGlySerPheArgLysG 293
 825 CCGCGCCGCGCAAAACACAGGCGCAACAGTGTCTTTCAGGAAAG 874
 293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
 875 GACTTGGAAATGTTGCCAGAAATATCTGCAAGATTAAGTGAACAAGTT 924
 310 LysLeuSerTyrLysLeuSerSerIleThrLysSerGluLysGlyGlyTyr 326
 925 AAGTGTCTGGAAAGCTCTCAGGTATCAGTAAAGCTGGAGAGCGGAGATA 974
 326 rHisLeuThrTyrGluThrProGluGlyValAlaIleArgAspGluArgL 343
 975 CAACCTTAACATTAAGAGATCCAGATGTTAGTTCCTGCGAGAGAAA 1024
 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
 1025 GTGTGTATTAAGGCGTCCATCTCAGATGTTGCAAGTGTCTGTCGCCCT 1074
 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
 1075 CTTTTCGAATCTGCTGCAATATGCACTCAAAAACATTAATACCCCACT 1124
 376 IGIValAlaIleThrIleSerTyrProGluAlaIleArgAspGluArgL 393
 1125 TGCAGCAGTATCTATCTCCTACCCCAAAAAGCAATCCAGACAGATGTT 1174
 393 euValAspGlyGluLeuLysGlyPheGlyGluLeuHisProArgThrGln 409
 1175 TGATFATGATGGTGAACAAAAGGGTGGTGGCAATTTGCACAGCGCA 1224
 410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
 1225 GGGATTGAACATTAAGAACTATCTACACCTCCTCCCTTCCAAATCG 1274
 426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyValAlaLysA 443
 1275 CGCACCGCCCGGAGAAATTTGGCTGTGAACATTAATGGCGGCTACAA 1324
 443 snProGluIleLeuSerLysTrpGluSerGlnLeuValGluValAlaAsp 459
 1325 ACACCGGAAATTTGCTCAGAGTGAAGTGAAGTGAAGCAGTGTGAC 1374
 460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
 1375 AGGATTTGAGGAAATTCCTAATTAAGCCCTAATTCAGCCAGATCCACTTAA 1424
 476 lValIleValArgValTrrProGluAlaIleProGlnPheLeuValGlyH 493
 1425 AATAGGAGTAAAGGTTATGGCCCAAGCCATTCCTCAGTTCAGTGGTGC 1474
 493 lIleuAspPheLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
 1475 ACTTTGATATTCCTTGACAGCGGCTAAATCATCTTAACGCTTCGGGCTTAC 1524
 510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
 1525 GAAGGCGCTATTTTGGGTCGCAATTAACGCTGCTGGTGAAGCTTAGGCGG 1574
 526 gcysValIleGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC47180
 seq_documentation_block:
 ID AAC47180 standard; DNA; 1690 BP.
 XX
 AC AAC47180;
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 52854.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 PD 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 PF
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135253.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136621.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137322.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.

Align seq 1/1 to: AAC47180 from: 1 to: 1690

22 LeuAlaIleuLeuSnaIrrhSerPheIleProPheSerSerIleSerIy 38
 7 TTAATCTCTCCGTCGACGACCAATTCGCTTCCGTCGCTTCCGAA 56
 38 ArgAsnSerValAsnCysAsnGlyTrp.....ArgThrArgCysS 52
 57 GCCCAAT...CTCCGATTAATGTTAAAGCCCTTTAGACTCCGTTGT 103
 52 eYValAlaIlyAspTYrTrValIProSerSerAlaValaSPglYglYPro 68
 104 CAGTGGCCGGTGGAGCACCGCTGGANCTTCAAAATTCAGAGCGGAGA 153
 69 AlaAlaIleuLeu.....AspCysValIleValGlyAlaIlyIleSerGI 83
 154 GGCACCCACATCACAGAGTGTGTGATTTGTCGGCGAGGTATFAGTGG 203
 83 YleuCYIleAlaIleValIleMetSerAlaAsnTYr.....ProA 96
 204 TCTTTGGATGCTCAGGCGCTGCTACTACTAAGCATCCCTGATGCTGCCGA 253
 96 snIleuMetValIThrGluAlaIArgAspArgAlaIlyIAsnIleThrThr 112
 254 ATTTAAATTTGGACCGAGCTAAGGATGCTGTGGAGGCAACATTAACACT 303
 113 ValGluArgAspGlyTYrTYrLeuTPgluGluGlyProAsnSerPheGlnPr 129
 304 CGTGAAGAAGATGTTTCTCTGCGGAAAGGTCCCAATFAGTTTCAACC 353
 129 oSerAspProMetIleuThrMetIleAlaValaIAspCysGlyLeuLYsAspSPl 146
 354 GTCTGATCTCATGCTCCTACTATGCTGAGTATGCTGTGAAAGGATGATT 403
 146 euValIleuGlyAspProAsnAlaIProArpPheValIleuTrpLYsGlyLYs 162
 404 TGGTGTGGAGAGATCTACTCCGCCCAAGTGTGTGTGGAGATGGGAAA 453
 163 LeuArgProValProSerLYsLeuThrAspLeuProPhePheAspLeuKe 179
 454 TTTGAGCGCGTTCCTCATTCAGACTTAACAGACTTACCGTCTTGTGATTGAT 503
 179 tSerIleProGlyLYsLeuArgAlaGlyPheGlyProIleGlyLeuArgP 196
 504 GAGTATTTGGTGGAGAGATTAAGACTGCTGTTGGTGCACCTTGGCATTGCAGC 553
 196 roSerProProGlyHisGluGluSerValGluGlnPheValaIArgArgAsn 212
 554 CGTCACTCCAGAGTCTGTGAAAGATCTGTGGAGGAGTGTGTACGGCGTAAAC 603
 213 LeuGlyGlyGluValPheGluArgLeuIleGluProPheCysSerGlyVa 229
 604 CTGCGATGATGAGGTTTGTAGCCCTGATTAACCGTTTGTCAAGTGT 653
 229 IYrValGlyAspProSerIlyLeuSerMetLYsAlaIlePheGlyLYsV 246
 654 TTTATGCTGGGATCTTCAAAACTGAGCATAAGCAAGCAGCTTTGGGAAGG 703
 246 alTrpLYsLeuGluGluIleuThrGlyIleSerIleIleGlyIleThrPheLYs 262
 704 TTTGGAAACTAGACAAATGATGAGAGCAATAATAGGTGATCTTTTAAAG 753
 263 AlaIleIleGlyGluArgSerSerThrProLYsAlaIleProArgAspProArqle 279
 754 GCAATFTCAGAGAGGAAAGAAAGCTCCCAAGGCAAGAGAGAGCGCGGCGCT 803
 279 uProLYsProLYsGlyGlnThyValGlySerPheArgLYsGlyLeuArgm 286
 804 GCCAAAACCCAGGCGCAAAACAGTGTGTTCTTTCAGAGAGGCGCTTCGAA 853
 296 deLeuProAspAlaIleSerAlaArgLeuGlySerLYsLeuLYsLeuSer 312
 854 TGTTCAGAGAGCAATATCTGCAAGATTAGGTAGCAAAAGTTAAGTTGTCT 903

313 TrpLYsLeuSerSerIleThrLYsSerGlyGlyTYrHisLeuTh 329
 904 TGGAAAGCTCAGGATACACTAAGCTGGAGAGCGGAGGATCAACTTAAC 953
 329 rTYrGluThrProGluGlyValValSerLeuGlnSerArgSerIleValIm 346
 954 ATATGAGACTCCAGATGTTTGTAGTTCCGTCAGAGCAAAAGTGTGTAA 1003
 346 eThrValProSerTYrValAlaSerAsnIleLeuArgProIleuSerVal 362
 1004 TCAAGCTGCATCTCATGTGGAAAGTGGCTCTTCCGCCCTTTTCTGAA 1053
 363 AlaAlaIleAspAlaLeuSerAsnPheTYrTYrProProValGlyAlaVal 379
 1054 TCTGCTGCAAAATGACTCAAAACTATATTAATTAACCAACAGTTCAGCAGT 1103
 379 IThrIleSerTYrProGluGluAlaIleArgAspGluArgLeuValaSPg 396
 1104 ATCTATCTGCTGACCGGAAAGCAATCCGAAACAGAAATGTTGATVAGATG 1153
 396 YgluIleuLYsGlyPheGlyGlnleuHisProArgThrGlnGlyValGlu 412
 1154 GTGAACCTAAGGCTTTGGCAATTTGCATCCACGCAACGAGGAGTTGAA 1203
 413 ThrLeuGlyThrIleTYrSerSerSerLeuPheProAsnArgAlaProLY 429
 1204 ACATTAAGCAACTATCTACAGCTCCTCACCTTCCAAATTCGCGCACCCGC 1253
 429 sGlyArgValleuLeuLeuAsnTYrIleGlyGlyAlaLYsAsnProGluI 446
 1254 CGGAAGAAATTTGGTGTGAACCTACATGTCGCGGCTCTCAACACACCGGAA 1303
 446 IeLeuSerLYsThrGluSerGlnleuValGluValaIAspArgAspLeu 462
 1304 TTTCTGTCCAAAGCTGAGGTGAGTTGAGAAAGCAAGTGTGACAGAGATTTG 1353
 463 ArgLYsMetLeuIleLYsPheLYsAlaGlnAspProLeuValValGlyVa 479
 1354 ACGAAATGCTTAATTAAGCTTAATTCGACCGCATCCACTTAATTAATFAGAGT 1403
 479 IArgValIrrProGlnAlaIleProGlnPheLeuValGlyHisLeuAspT 496
 1404 TTAGGATATGGCTCAAGCCATTCCTCAGTTCCTACTGTGTGCACTTTGATA 1453
 496 hrLeuSerThrAlaLYsAlaIleMetAsnAspAsnGlyLeuGluGlyLeu 512
 1454 TCCCTTGACAGCGCTAATATCATCTCTTAACGCTCTTCCGGCTACGAAAGGCTTA 1503
 513 PheLeuGlyGlyAsnTYrValSerGlyValAlaIleLeuGlyArgCysValGI 529
 1504 TTTTGGGGGGCAATTAAGCTGCTGTGTGAGCCCTTAAGCCCGGTGTGAGA 1553
 529 uGlyAlaIleTYrGluValAlaIleSerGlyValIleThrGlyPheLeuSerArgTYrA 546
 1554 AGGCGCATATGTAACCCGCGATTAAGGATCAACAACTTCAACAGCGTACG 1603
 546 IatTYrLYs 548
 1604 CTTTACAG 1611
 seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76571
 seq_documentation_block:
 ID AAF76571 standard; cdNA; 1719 BP.
 XX AAF76571;
 AC
 XX 16-MAY-2001 (first entry)
 DT
 XX Arabidopsis protoporphyrinogen oxidase coding sequence SEQ ID NO: 1.
 DE
 XX Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;


```

166 aLProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 TGCCGGGAACCTGACTGATTTGGCTTTTGGACTTGATGAGCAATGGT 531
183 GlyLysLeuAraGalaGlyPheGlyProIleGlyLeuAraGProSerPro 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
532 GGCAAAATCAGAGGGCGGCTTGGTGGCTTGGAAATGGCCCTCCCTCC 581
199 OGLYHISGLUGLUSERVALGLUINPHEVALAARGASMLUUGLYG 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
582 AGGTCATGAGCAATCGGTTGAAAGAGTTGTCGCGGAACTTGGTANTG 631
216 LuvalPheGluAraGluIleGluProPheCysSerGlyValIleValGly 232
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
632 AGGTTTGAACGGTTGATAGAGCCCTTTTGTCAAGGGGCTATAGCGAGC 681
233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValIleLys 249
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
682 GATCCTCAAAATTAAGTAAGAAAGCAGCATTCGGGAAAGTTGGAAGCT 731
249 uGIUGLUTHrGLYGLYSerIleIleGlyGLYThPheLysAlaIleLysG 266
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
732 GGAAAAAATGGTGGTACCATTTAGTGGGAACCTTCAAAACCAATACAG 781
266 LuAraSerSerThrProLysAlaProAraGAspProAraGLeuProLysPro 282
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
782 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCCGCTCCCAAAACA 831
283 LysGlyGluThrValGlySerPheAraGlyLysLeuAraGMetLeuProAs 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
832 AAAGGTCAGACTGTTGATCTTTCGGGAAGGACTTACCATGTTGCCCTGA 881
299 PalaIleSerAlaAraGLeuGlySerLysLeuLysSerTrpLysLeus 316
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
882 TGCAAATTCGCCAGACATAGCAACAAAGTAAGTTATCTTGGAAAGCTTT 931
316 eSerIleThrLysSerGlyLysGlyGlyGlyThrHisLysLeuThrTYRGLUThr 332
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
932 CAAGTATTAGTAACACTGGATAGTGAGAGTAGACAGTTGACATATGAAACA 981
333 ProGluGlyValValSerLeuGlnSerAraGSerIleValMetThrValPr 349
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
982 CCAGAAGGAGTGGTTCCTTTCAGTGGCAAAACTGTTCTCCAGCCATTTCC 1031
349 oSerTYrValAlaSerAsnIleLeuAraGProLeuSerValAlaAlaAla 366
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1032 TTCCTATGTGGTACTAGTACATGTCGCGTCTGCTGCTGCTGCGAG 1081
366 sPAlaIleSerAsnPheTYrTYrProValGlyAlaValThrIleSer 382
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1082 ATGCACCTTCAAGAGTTTATTAACCTCCAGTTGCTGCGAGTTCCATATCC 1131
383 TYRProGluAlaAlaIleAraGAspGluAraGLeuValAspGlyGluLeuLys 399
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1132 TATCCAAAAGAGCTAATAGATCAGAAATGCTTATGATGATGGTGA 1181
399 sGlyLysPheGlyGluLeuHisProAraGThrGlnGlyValGluThrLeuGlyT 416
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1182 GGGGTTTGGTCAATTCATCCACGATGCAAGGAGGAAATCATTAGGAA 1231
416 hrIleTYrSerSerSerLeuPheProAsnAraGAlaProLysGlyAraGVal 432
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1232 CATATATACAGCTCATCTATTTCCCAACCGACGACACACCGGAAGGTT 1281
433 LeuLeuLeuAsnTYrIleGlyGlyAlaLysAsnProGluIleLeuSerLys 449
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1282 CTACTCTTGAATTAACATTCATGGAGGACCAACTAATCTGGAATTTATCGAA 1331
449 sThrcLysSerGlnLeuValGluValValAspAraGAspLeuAraGlyMetL 466
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1332 GACGGACAGTGAACCTTGGGAAACAGTGTGATCGAGATTTGGAGAAATCC 1381

```

```

466 eutIleYsProLysAlaGlnAspProLeuValValGlyValAraGValTrp 482
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1382 TTRATACCACCAATGCCAGATCCATTTGTAGTGGGGGTGAGACTGGT 1431
483 ProGluAlaIleProGlnPheLeuValGlyHisLysAspThrLeuSerTh 499
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1432 CCTCAAGCTATTTCCACAGTTCATGTTGGCCAMCTTGATCTTGATGT 1481
499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGly 516
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1482 TGCTAAAGCTTCTATACGAAATACGTGGTTTGAAGGCTCTTCTGGGG 1531
516 lYAsnTYrValSerGlyValAlaLeuGlyAraGcysValGluValAraYr 532
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1532 GTAATTTATGTGTGGTGGCTTGGGAGCATCCGCTTGAGGGAGGCTAT 1581
533 GluValAlaSerGluValThrGlyPheLeuSerAraGTYrAlaTYrLys 548
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1582 GAGGTAGCAGCTGAGTAAACGATTTTCTCACAAAPAGAGTGTACAAA 1629
seq_name: /SID58/gcgdata/geneseq/geneseqn/NA1197.DAT:AAV04309
seq_documentation_block:
ID AAV04309 standard; cDNA; 1847 BP.
XX
AC AAV04309;
XX
DT 20-APR-1998 (first entry)
XX
DE Soybean protox-1 cDNA.
XX
KW Protoporphyrirogen oxidase-1; protox-1; promoter; soybean;
KW herbicide resistance; breeding programme; probe; gene isolation;
KW genomic mapping; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 55..1686
FT FT /*tag= a
FT FT /product= protox-1
XX
PN W09732028-A1.
XX
PD 04-SEP-1997.
XX
PE 27-FEB-1997; 97WO-US03343.
XX
PR 21-JUN-1996; 96US-0020003.
PR 28-FEB-1996; 96US-0012705.
PR 28-FEB-1996; 96US-0013612.
XX
PA (NOVS ) NOVARTIS AG.
XX
PI Johnson MA, Volrath SL, Ward ER:
XX
DR MPI: 1997-489209/45.
XX
DR P-PSDB: AAW41608.
XX
PT DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT optionally linked to a heterologous gene, especially to express
PT herbicide-resistant enzymes, and plants containing such constructs
XX
PS Disclosure: Pages 68-71; 114pp; English.
XX
CC The present sequence encodes soybean protoporphyrirogen oxidase-1
CC (protox-1).
CC The protox-1 promoter can be used to express herbicide resistant
CC enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC containing a chimeric gene of the promoter and a heterologous
CC coding sequence. The plant can also be used in breeding programmes.
CC Also hybridising fragments of the protox coding sequence can be
CC used as probes, e.g. to isolate related genes or for genomic

```

CC mapping.
XX Sequence 1847 BP; 473 A; 416 C; 453 G; 505 T; 0 other:

alignment_scores:
Quality: 2062.50 Length: 550
Ratio: 4.218 Gaps: 7
Percent Similarity: 88.909 Percent Identity: 73.273

alignment_block:
US-09-508-418-2 x AAV04309
Align seg 1/1 to: AAV04309 from: 1 to: 1847

3 ThrThrProIleAlaIAsnHisProAsnIlePheThrHisGlnSerSer 19
102 TCCTCCGCCCTCCCTCCATTCGCCCAACCTCTTTCACCTCCACACTC 151
19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSer 36
152 GAAA.....ATTCCC.....T 162
36 Le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
163 CGCTCTCCCTTAACCTATP.....CTACGCTGCTC 194
52 ValAlaIAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
195 CATTCCGGAGAGATCCACCGCGTCTCCGCCAAAACACAGACACTCCGCC 244
68 roAlaIAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
245 GC.....GTGACCTGGCTGCTGCTGGCGAGGCGCTCACAGCCCTC 285
85 CysIleAlaGlnValMetSerAlaAsnTyr.....ProAsnLeuMetVal 99
286 TGGATTCGCCACCGCCCTCCGCCACCAACACGCCAATGCCAACGCTCCTG 335
99 IThrGluAlaIAlaArgAspArgIaGlyIAsnIleThrThrValIAlaArgA 116
336 CACGGAGGCCCGAGACCGCGTCCGGCGCAACATCACACAGATGAGARAGG 385
116 spGlyTyrLeuTrpGluGlyIProAsnSerPheGlnProSerAspPro 132
386 ACCGATACCTCTGGGAAAGAGGCCCAACAGCTTCCAGCCCTTCTGTACCA 435
133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuG 149
436 ATGCTGCACCAATGGGTGTGGACAGTGTAAAGATGAGAGCTGTGGGG 485
149 YAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
486 GGATCCCTGATGCACCTCGGTTGTGTGGTGGAAACAGAGTTGAGGCCGG 535
166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
536 TGCCTCCGGGAAAGCTGACTGATTTCCCTTCTTGAATTCATTTGAGCA 585
183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
586 GGGAAAATTCAGGGCTGGCTTGGTGGCGCTTGAATTCGGCTCCCTCC 635
199 oGlyHisGlnGlnSerValGlnGlnPheValArgArgAsnLeuGlyGly 216
636 AGGTCAATGAGGAGATCGGTTGAAAGATTTGTTCGTCGGAACCTTGGGAG 685
216 LuValPheGluArgLeuIleGlnProPheCysSerGlyValIAlaTyr 232
686 AGGTTTTTGAACGGTTGATAGACCTTTTGTTCAGGGGTCTATGACAGGC 735
233 AspProSerLysLeuSerMetLysAlaIAlaPheGlyLysValTrpLysLe 249

736 GATCCTTCAAAAATTAAGTATGAAAGCACAGACTTCGGGAAAAGTTTGGAACT 785
249 uGluGluThrIhrGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
786 GGAATAAATAATGGTGTAGCATTTATTTGGTGAACCTTTCACAGCAATACAG 835
266 LuArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
836 AGAGAATAAGGAGCTTCAAAACACCTCGAAGATCCGCGTCCGCAAAACCA 885
283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
886 AAAGGTACAGACTGTGATCTTCCGGAAGGACTTACCACTGATGCTCCGA 935
299 PalIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
936 TGCAAATTTCTGCCACAGACTRAGCAACAAGTAAAGTTATCTTGGAGACTT 985
316 erSerIleThrLysSerGlyLysGlyIYrHisLeuThrTrpGluThr 332
986 CAAGTATTAGTAACTGGATAGTAGAGAGTACAGTTGACATPAGAAACA 1035
333 ProGluGlyValIAlaSerLeuGlnSerArgSerIleValMetThrValPr 349
1036 CCAGAAGGAGTGTCTTTCGACGCAAAACTGTGCTCAGACCACTTCC 1085
349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaIAla 366
1086 TTCATATGTGTAGTACATTCCTGCTCCTGCTGCTGCTGCTGCTGCTG 1135
366 spAlaLeuSerAsnPheTyrTrpProValGlyAlaValIAlaThrIleSer 382
1136 ATGCACCTTCAAAAGTTTATTTACCTCCACAGTTCGCTGCTGCTGCTG 1185
383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuL 399
1186 TATCCAAAAGAGCTTATGATGACAGAAATGCTGATGATGATGATGATG 1235
399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
1236 GGGGTTTGGTCAATTCATCCACGTAGCCAAAGAGTGGAAAACATTTAGGA 1285
416 hIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
1286 CTATPATAAGCTCATCACTATTCCTCCCAACCGAGCCACCTGGAAGGG 1335
433 LeuLeuLeuAsnTyrIleGlyIAlaLysAsnProGluIleLeuSerL 449
1336 CTACTCTGAAATTTACATTTGGAGAGCAACTAATACCTGGAATTTATTCG 1385
449 sThrGlnSerGlnLeuValGluValIAlaAspArgAspLeuArgLysMet 466
1386 GACGGACAGTGAACCTTGTGGAAACAGTGTGATCGAGATTTGAGGAAAAACC 1435
466 euIleLysProLysAlaGlnAspProLeuValIAlaGlyValArgValTrp 482
1436 TTAATAAACCCAAATGCCAGGATCCATTTGTAGTGGGGGTGAGACTGTG 1485
483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspTrpIleLeuSer 499
1486 CCTCAAGCTATTTCCACAGTCTTAACTTGGCCATCTTGAATCTTGTAGAG 1535
499 rAlaLysAlaIAlaMetAsnAspAsnGlyLeuGlnGlyLeuPheLeuGly 516
1536 TGCTAAAGCTTCTATACAGAAATACGAGGTTTGAAGGGGCTCTCCCTGG 1585
516 LysAsnTyrValSerGlyValAlaLeuGlyArgCysValGlnGlyAlaTyr 532
1586 GTAATAATGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1635
533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
1636 GAGGTAGCGCTGAGAGTAAACGATTTCTCAAAAATTAAGAGTACAAA 1683

seq_name: /SIDS8/gcgdata/geneseq/geneseq/NA1997.DAT:AA186122

seq_documentation_block: ID: AA186122 standard; CDNA: 1847 BP.

AA186122; 01-MAR-1998 (first entry) Soybean protoporphyrinogen oxidase (protox-1) cDNA clone pMDC-12. Protox-1; protoporphyrinogen oxidase; inhibitor; soybean; herbicide tolerance; herbicide resistance; transgenic plant; ss. Glycine max var. Williams 82. Location/Qualifiers Key: 55..1686 /*tag= a M09732011-A1. 04-SEP-1997. 27-FEB-1997; 97WO-US03313. 21-JUN-1996; 96US-0020003. 28-FEB-1996; 96US-0012705. 28-FEB-1996; 96US-0013612. (NOVS) NOVARTIS AG. Helfetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER. MPI: 1997-448683/41. P-PSDB: AAM25739. New DNA encoding plant protoporphyrinogen oxidase enzyme - and herbicide resistant mutants, useful to prepare plants resistant to herbicide which therefore kills undesired vegetation only. Claim 6; Page 127-130; 196pp; English. This cDNA clone codes for soybean protoporphyrinogen oxidase (Protox-1, see AAM25739). It was isolated from an epicotyl cDNA library using Arabidopsis Protox-1 cDNA (see AA186129) as probe. Soybean Protox-1 in plasmid vector SK vector has been deposited as pMDC-12 (NRRL AAB21516). The isolated cDNA can be mutated so that it encodes a modified protox enzyme that is resistant to protox inhibitors and hence is herbicide tolerant. Plants, especially crop plants, may be engineered for resistance to protox inhibitors via mutation of the native protox gene to a resistant form, or they may be transformed with a gene encoding an inhibitor-resistant form of a plant protox enzyme, including claimed forms from wheat, soybean, cotton, sugarbeet, oilseed rape, rice and sorghum (see AAM25738-45). Application of herbicide therefore kills undesired vegetation only. Plant protox nucleic acids can also be used as probes and PCR primers, as selectable markers in plant cell transformation methods, and for recombinant production of protox enzymes in host cells. Sequence 1847 BP; 473 A; 416 C; 453 G; 505 T; 0 other;

Alignment_scores: Quality: 2062.50 Length: 550 Ratio: 4.218 Gaps: 7 Percent Similarity: 88.909 Percent Identity: 73.273

alignment_block: US-09-508-418-2 x AA186122

Align seg 1/1 to: AA186122 from: 1 to: 1847

3 ThrThrProTlleAlaAsnHisProAsnIlePheThrHisGlnSerSer 19
102 TCTTGGCCCTCCCTCCATTCCCAACCTTTCTTCTTCCCTCCACATC 151
19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSer 36
152 GAAA.....ATATCC.....T 162
36 Le.SerIysArgAsnSerValAsnCysAsnGlyTyrPargThrArgCys 52
163 CGCTCTGGCCCTAACCTATT.....CTACGCTGGCTC 194
52 rValAlaIysAspTyrThrVal...ProSerSerAlaValAspGlyGly 68
195 CATTGGGAGAGATCCACCGCTCTCCGCCCAAAACAGAGACTCCGCC 244
68 rAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
245 CC.....GTGGACTGGCTGGTGGGGAGGCGGTACAGCGGCTC 285
85 CysTlleAlaGlnValMetSerAlaAsnTyr.....ProAsnLeuMetVal 99
286 TGCATGCCCCAGGCCCTCCGCCCAAAACAGCCCAATGCCAATGCCGTC 335
99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
336 CACGAGGCCCGAGACCGCTGGCGGCAACATCACCATGAGAAAGG 385
116 spGlyTyrLeuTrrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
386 ACGGATACCTCTGGGAAGAGGCCCAACAGCTTCCAGCTTGTGATCCA 435
133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuG 149
436 ATGCTCACCAATGGTGTGGACAGTGTAAAGAGATGCTTGTGGTGG 485
149 yAspProAsnAlaProArgPheValLeuTrrpLysGlyLysLeuArgPro 166
486 GGATCTGATGACCTCTGGTGTGTGGTGGTGGTGGTGGTGGTGGTGG 535
166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
536 TCGCCGGGAAAGCTGACTGATTTGGCTTTTGGACTTGAAGACATTTGG 585
183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerPro 199
586 GGCMAAATCAAGGGCTGGCTTTGGTGGCTTTGGTGGTGGTGGTGGTGG 635
199 oGlyHisGluGlnSerValGluGlnPheValArgArgAsnLeuGlyGly 216
636 AGGTCATGAGGAATGGTTGAAGGTTTGGTTCGTCGGAACCTTGGTGAAG 685
216 lValIlePheGluArgLeuIleGluProPheCysSerGlyValYrValGly 232
686 ACGTTTTTGAACGGTTGATGAGCCCTTTTGTTCAGGGGCTATGAGAGC 735
233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrrpLysLe 249
736 GATCCCTTCAAAATTAAGTATGAAGAAGCAATTTGGGAAAGTTTGGAA 785
249 uGluGluThrGlySerIleIleGlyGlyGlyThrPheLysAlaIleLysG 266
786 GGAATAAATAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 835
266 lArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
836 AAGAATAATGAGCTTCAAAACCACTCGAGATCCGGGTGGCCAAACCA 885
283 LysGlyGlnThrValGlySerPheArgGlyGlyLeuArgMetLeuProAs 299
886 AAAGTCAAGACTGTGGATCTTCCGGAAGGAGCACTTACCAATGTTGGCTGA 935

```

299 PALAlleSeRAlAlArLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
    |||||||.....:..||| |
936 TCGAAATTCTCCAGACTAGGCAACAAGTAAGTTCCTGGAAAGCGTTT 985
316 eRSerLeThrSerGlyLysGlyGlyGlyGlyThrHisLeuThrThyGluThr 332
    |||||||.....:..||| |
966 CAAGTATTTAGTAAACCTGGATAGTGGAGATACAGTTCGATGATAGGAACA 1035
333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetPhrValPr 349
    |||||||.....:..||| |
1036 CCAGAAGAGAGGTTCCTTCAGTGGAAAAAAGTTGCTCCGACCATTCC 1085
349 oSeTTrValAlAAserAsnIleLeuArgProLeuSerValAlAlAlAlAlA 366
    |||||||.....:..||| |
1086 TTCCTAGTTCCTGATACATTGCTGCCCTCCTGTCCTGCTGCTGCAG 1135
366 sPAlAlleSeRAsnPhetTrpTrpProProValAlValAlAlAlAlAlSer 382
    |||||||.....:..||| |
1136 ATGCACCTTCAAAGTTTATTATCCCTCCAGTGGCTGCCAATTCATATTC 1185
383 TyrProGlnGluAlAlAlAArgAspGluArgLeuValAspGlyGluLeuLy 399
    |||||||.....:..||| |
1186 TATCCAAAAGAGCTATTAGATCGAATGCTGATAGATGGTGAGTTGAA 1235
399 sGlyPhgGlyGlyLeuHisProArgTrpGlnGlyValGluTrpLeuGlyT 416
    |||||||.....:..||| |
1236 GGGGTTTGGTCAAATTCATCCATCCATCCAAAGAGGTGAAACATTAGAAA 1285
416 hrIleTyrSerSerLeuPheProAsnArgAlAProLysGlyArgVal 432
    |||||||.....:..||| |
1286 CTAAATACAGTCACTACTATTCCTCCCAACAGACCACTGGAGGAGT 1335
433 LeuLeuLeuAsnTyrIleGlyGlyAlAlLysAsnProGlnIleLeuSerLy 449
    |||||||.....:..||| |
1336 CTACTCTTGAATTCATTTGGAGAGGCAACTAATACCTGGAAATTTAATCGAA 1385
449 sThrGlnSerGlnLeuValGluValAlAspArgAspLeuArgLysMetL 466
    |||||||.....:..||| |
1386 GACCGACAGCTGAACTGTGGAAACAGTGTGATCGAGATTTGAGGAAAATCC 1435
466 euLeLysProLysAlAlAlAspProLeuValAlGlyValAlArgValTrp 482
    |||||||.....:..||| |
1436 TTATAAACCCAAATGGCCAGATCCATTTGATGGGGGAGACACTGTGG 1485
483 ProGlnAlAlLeProGlnPheLeuValAlGlyHisLeuAspThrLeuSerTh 499
    |||||||.....:..||| |
1486 CCTCAACCTATTCACAGCTTCAGTGGCCATCTGATCTTCCAGATCT 1535
499 rAlAlYsAlAlAlMeLAsnAspAsnGlyLeuGlyGlyLeuPheLeuGlyG 516
    |||||||.....:..||| |
1536 TGCCTAAAGCTTTATTCAGAAATCTGGGTTTTGAAGGCGCTTCCTGGGG 1585
516 LysAsnTyrValSerGlyValAlAlLeuGlyArgCysValAlGluGlyAlaTyr 532
    |||||||.....:..||| |
1586 GTAATATATTGTCTGGGTGCTGCCTGGGAGATGGCTGTGAGGAGACCTAN 1635
533 GluValAlAlAserGluValAlThrGlyPheLeuSerArgTyrAlaTyrLys 548
    |||||||.....:..||| |
1636 GAGGTAGCAGCTGAAAGTAAACGATTTTCTACAAATAGAGTAAACAA 1683
seq_name: /SID8/gcgsdata/geneseq/geneseq/nA2001.DAT:AAF76576

```

```

seq_documentation_block:
ID   AAF76576 standard; cDNA; 1847 bp.
XX   AAF76576;
XX
XX   16-MAY-2001 (first entry)
DE   Soybean protoporphyrinogen oxidase coding sequence SEQ ID NO: 11.
XX   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW

```

```

KW soybean; sugar beet; oilseed rape; sugar cane; mutant; mutlein; ss.
XX Glycine max.
OS
XX WO200112825-A1.
XX
XX 22-FEB-2001.
XX
XX 30-JUN-2000; 2000WO-EP06127.
XX
XX 13-AUG-1999; 99US-0373691.
XX
XX (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX Johnson MA, Volrath SL, Helfetz PB, Law MD;
XX WPI: 2001-234914/24.
XX P-PSDB; AAB72906.
XX
XX Plant DNA molecules encoding herbicide-tolerant forms of
XX protoporphyrinogen oxidase which are useful for rationally designing
XX new inhibitory herbicides and for producing herbicide-tolerant
XX transgenic plants and seeds -
XX Claim 2; Page 168-172; 228pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
XX sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
XX oxidase (protoporphyrinogen oxidase). Examples of these mutants are shown in
XX AAB72906-AAB72926. They are useful as they enable the production of
XX herbicide-tolerant plants and seeds. The present sequence is a protox
XX coding sequence.
XX
XX Sequence 1847 bp; 473 A; 416 C; 453 G; 505 T; 0 other;

```

```

alignment_scores:
Quality: 2062.50    Length: 550
Ratio: 4.218       Gaps: 7
Percent Similarity: 88.909   Percent Identity: 73.273

```

```

alignment_block:
US-09-508-418-2 x AAF76576 ..
Align seg 1/1 to: AAF76576 from: 1 to: 1847
3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
  |||||.....:..||| |
102 FCTTGGCCCCCTCCATTCCTCCCAACCTCTTCTGCTCACCTCCACTCC 151
19 rSerProLeuAlAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
  |||||
152 GAAA.....:..||| |
36 Le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
  |||||.....:..||| |
163 CCTCTCGGCCCTAACCCATAAT.....:..CTACGGCTGCTC 194
52 rValAlAlalAspAspTyrThrVal...ProSerSerAlAlValAspGlyGlyP 68
  |||||.....:..||| |
195 CAYTCTGGAGGAAATCCACCCTTCCTCCGCCCAAAAGCAGACACTCCGCC 244
68 rGAlAlAlAlGluLeuAspCysValIleValAlGlyAlaGlyIleSerGlyLeu 84
  |||||.....:..||| |
245 CC.....:..GTGACTGCCTGCTCCTCGGCGGAGGAGGCTGACAGCGCC 285
85 CysIleAlAlGluValMetSerAlaAsnTyr.....ProAsnLeuMetVal 99
  |||||.....:..||| |
286 TGCATCCGCCAGGCCCTCCGCAACCAACCGCAATGCCAATGCCATGCTGCT 335
99 lThrGlnAlAlArGAspArgAlAlGlyGlyAsnIleThrThrValAlGluArg 116
  |||||.....:..||| |

```

336 CACGGAGGCCCGAAGACCGCGTGGCGGACATGACCAACGATGGAGAGCG 385
116 SPGLYTRLeuTRpGLUGLUGLIPROAsnSerPheGlnProSerAspPro 132
386 ACGGATNCCCTGGGAAGAAAGCCCAACAGCTTCCAGCCCTTCGATCCA 435
133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuG 149
436 ATGCTCACCATGTGTGTGGACAGTGTAAAGATGATGAGCTTGTGGGG 485
149 YASPRoKsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
486 GGATCCGATGCACCTCGGTTGTGTGTGGAAACAGAGAGTTGGAGCCGG 535
166 aLProSerLysLeuThrAspLeuProPhePheAspLeuMetSerLe 182
536 TGCCCGGAGACTGACTGATTTGGCTTCTTTCGACTTGATGAGCATGGT 585
183 GLYLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerPro 199
586 GGCAAAAATCAGGGCTGTGGCTTGTGGCCGCTTGGAAATTCGGCCCTCC 635
199 cGLYHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGly 216
636 AGGTCATGAGGAAATCGTTGAAGAGTTTGTTCGCGGAACTTGGTGTG 685
216 LuValPheGluArgLeuIleGluProPheCysSerGlyValTyrValG 232
686 AGGTTTTGAACGGTTGATAGAGCCCTTTTGTTCAGGGGCTCATNGCAG 735
233 ASPPROSerLysLeuSerMetLysAlaIlePheGlyLysValTyrLys 249
736 GATCCCTGAAAATTAAGATGAAAGACAGCATTCGGGAAAATTTGGAAAG 785
249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLys 266
786 GGAATAAATATGGTGTAGCATATATTTGGTGAACCTTCAACCAATATCA 835
266 LuArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
836 AGAGAAAATGGACCTTCAAAAACCCTCGAGATCCCGCTGCCCAAAACCA 885
283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuPro 299
886 AAAGGTACAGACTGTGGATCTTCCGGAAGGACTTACCAGTTCCTGCA 935
299 PAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTyrLysLe 316
936 TCCAAATTCCTGCCAGACTAGCCAAACAAGTAAAGTATATCTTGGAAAG 985
316 eTSerIleThrLysSerGluLysGlyTyrHisIleLeuThrTyrGlu 332
986 CAAGTATTAAGTAACSTGATAGTGAAGATACAGTTTGCATATGAAACA 1035
333 PROGLUGLYValValSerLeuGlnSerArgSerIleValMetThrVal 349
1036 CCAGAAAGAGTGTCTTTCGACAGTGCAAAACCTGTTCGCTGACCATTC 1085
349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAla 366
1086 TTCCTATGTTCCTAGTACATGTCTGCTGCTGCTGCTGCTGCTGCTG 1135
366 sPAlaLeuSerAsnPheTyrTyrProProValGlyAlaValAlaThrL 382
1136 ATGCACTTCAAAAGTTTATATACCTCCAGTGTGCGAGTTTCCAAATCC 1185
383 TYRPROGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeu 399
1186 TATCCAAAAGACTATTTGATCAGAAATGCTGATAGATGATGAGTGA 1235
399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGly 416
1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAAGAGTGGAAACATTAAG 1285

416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
1286 CFATATACAGCTCATATCCATATCCCAACAGCAGCACACCTGGAGGGTT 1335
433 LeuLeuAsnTyrTrIleGlyAlaLysAsnProGluIleLeuSerLys 449
1336 CTACTCTGAAATTAATTTGGAGGAGCACTAAATFACTGGAAATTTATCCA 1385
449 sThrGluSerGlnLeuValGluValAlaAspArgAspLeuArgLysMet 466
1386 GACGACAGTGAACCTGTGGAAAACAGTGTGATCGAGATTTGAGGAAAATCC 1435
466 euIleLysProLysAlaGlnAspProLeuValGlyValArgValTyr 482
1436 TTTTAAACCCTTCAATGCCAGGATCCATTTGTGTGTGGGGGTGAGACTGTG 1485
483 PROGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSer 499
1486 CTTCAAGCTTATTCACAGCTTCTTGTGGCCATCTTGATCTGTAGATGT 1535
499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGly 516
1536 TGCATAAAGCTTCTATCAGAAATCTGTGGTGGAAAGGCTTCCCTTGGGG 1585
516 LysAsnTyrValSerGlyValAlaLeuGlyValGlyValGlyAlaTyr 532
1586 GTAATTAATGTCTGTGGTTCCTTGGAGAGATGCCCTTGAAGGAGGCTTAT 1635
533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
1636 GAGGTGACAGCTGAGTAAACGATTTTCTCACAAAATATAGAGTGAACAA 1683

OM of: US-09-508-418-2 to: Issued_Patents_NA:* out_format : pfs
Date: Jul 3, 2001 11:38 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame_psn.model -DEV=xip
-Q/cgn2_1/USPTO_spool/US09508418/runat_03072001_100017_24452/app-query.fasta_1.613
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.500
-GAPOP=4.500 -OGAPEXT=0.050 -XGAPP=10.000 -XGAPEXT=0.500
-PELOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09508418@cgn1_1.66 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLIPX
-WAIT -THREADS=1

Search information block:

Query: US-09-508-418-2
Query Length: 548
Database: Issued_Patents_NA:*
Database sequences: 317530
Search time (sec): 94.000000

score_list:

Table with columns: Sequence, Strd Orig, ZScore, BScore, Len, Documentation. Contains multiple rows of sequence identifiers and their corresponding scores and lengths.

/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-472-028A-3 + 490.50 910.02 4.4e-43 1738
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-3 + 490.50 910.02 4.4e-43 1738
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-3 + 490.50 910.02 4.4e-43 1738
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-3 + 490.50 910.02 4.4e-43 1738
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-3 + 490.50 910.02 4.4e-43 1738

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-15

seq documentation block:

Sequence 15 Application US/08808931
Patent No. 5939602
GENERAL INFORMATION:
APPLICANT: Voltath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Polter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Protoopryrinogen Oxidase and Inhibitor-Resistant Mutants
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESS: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,931
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1826 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum (cotton)
IMMEDIATE SOURCE:
CLONE: pMDC-15 (NRRL B-21594)
FEATURES:
NAME/KEY: misc_feature
LOCATION: 31..1647
OTHER INFORMATION: /product= "Cotton protox-1 coding
OTHER INFORMATION: region"
US-08-808-931-15

Alignment_scores:
Quality: 2113.00 Length: 506
Ratio: 4.477 Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249

alignment_block:
US-09-508-418-2 x US-08-808-931-15

Align seq 1/1 to: US-08-808-931-15 From: 1 to: 1826

47 TrrprgrhrtrgCysSerValAlaLysApsrYrthValProSerSerAl 63
130 TTCAGCTCCAGTCCCTCCGCGGACCCACGATTTCCCATCTAA 179
63 aValAspGlyProAlaAlaGluLeuAspCysValIleValGlyLaG 80
180 AATGAGGGGGGAGAATCCATCGGGGATTCGCATCGTTGGAGGTG 229
80 lylIeSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr... 94
230 GTATCAGAGGACTTGTGATTCATCCAGCTCCGCCACCAAGCCGTGAC 279
95 ProsnLeuMetValThrGluAlaArgAspArgAlaGlyLys 109
280 GTCCGTTCCAAATGTGATTTGTGACGGAGCCGACACCCTGTGGGCA 329
109 nIlethrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsn 126
330 CATCCATCCGTTGAGAGATGATATCTGTGGGAAGAAAGCCCAACA 379
126 erpHeGlnProSerAspProkMetLeuThrMetAlaValAspCysGlyLeu 142
380 GTTTCACCCCTCGATCTATTTCAACCATGCGCGGATAGTGATTCG 429
143 LysAspAspLeuValIleuGlyAspProAsnAlaProArpPheValIleuTr 159
430 AAGGACGATTTGGTTTGTGACCCCTAATGACCCGATTTGTACATAG 479
159 pLysGlyLysLeuArqProValProSerLysLeuThrAspLeuProPheP 176
480 GGAGGGAAACTAAGAGCCCTGCCTCCCAAGCAACCGACTGGCGCTTT 529
176 heAspLeuMetSerIleProGlyLysLeuArqAlaGlyPheGlyProIle 192
530 TTGATTTGATGACATTTGCTGGAATACTTAGGGCTGGGCTTCTG 579
193 GlyLeuArqProSerProProGlyHisGluGluSerValIleGluGlnPheVa 209
580 GGCATTCCGGCCCTCCCTCCGGTTATGAAGAAATCGGTGGAGGATTTGT 629
209 lArqArgAsnLeuGlyGluValPheGluArqLeuIleGluProPheC 226
630 GCGCCGTAATCTTGTGTGAGGTTTGTGAAAGCCTTATGTACCATTTT 679
226 ySeSerGlyValIleValAlaGlyAspProSerLysLeuSerMetLysAlaAla 242
680 GTTCAGCTTCTTATGACGGAGTCCCTTCAAAATTAAGCATGAACAACACA 729
243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyG 259
730 TTTGGAAAGGTATGGAAAGCTAGAAAGATGTTGCTGCACACATCATTTGTGG 779
259 YrthrPheLysAlaIleLysGluArqSerSerThrProLysAlaProArqA 276
780 CACTTTCAGACAAATCCAGGAGAAATAAGACACCTAAGCCACCCCAAG 829
276 sPProArqLeuProLysProLysGlyGlnThrValGlySerPheArqLys 292
830 ACCCGGCTGTGCCAAAACCGAAGGGCCAAACACGTTGGATCTTTTGAAG 879
293 GlyLeuArqMetLeuProAspAlaIleSerAlaArqLeuGlySerLysLe 309

880 GGACCTTACCATGCTGCCCTGAGGCAATTCGTAAACAGTWTGGTAGCAATGT 929
309 ulLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyL 326
930 AAAATATCTTGGAAGCTTCCAGATTAACCAAAATTGGGCATGGAGG 979
326 YrHisLeuThrTyrGluTrpProGluGlyValIleSerLeuGlnSerArq 342
980 ATAACTTGACATTTGAAACACCTGAAAGAAATGATCTTCAGAGATGA 1029
343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArqPr 359
1030 AGTGTGTAAATGACCATTCCATCCATGTCACAGTAACTTGTTCATCC 1079
359 OleuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTrpProPov 376
1080 TCTCTCCGGTGGCTGCGGACGATGCATTTCCCAATTTTATTTTCTCCAG 1129
376 aIGlyValAlaValThrIleSerTyrProGlnGluAlaIleArqAspGluArq 392
1130 TTGCATCAGTCCACAGTCTCCTAATCCAAAGAACCCATTCGAAAGAAATGT 1179
393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArqTrpG 409
1180 TTGATTTGATGTGMACTTAAGGGGTTGGCCAGTTGCACCCGACGCCA 1229
409 nGlyValGluThrLeuGlyThrIleTyrSerSerLeuPheProAsnA 426
1230 AGCAATTTGAAACTTTTGGACGATATACAGTTTCATCCATCTTTCCCAATTC 1279
426 rGAlaProLysGlyArqValIleLeuLeuAsnTyrIleGlyGlyAlaLys 442
1280 GAGCTCCATCTGCGACGGGTGTGCTGTGAACCTACTAAGAGAGAGCTACC 1329
443 AsnProGluIleLeuSerLysTrpGluSerGlnLeuValGluValAlas 459
1330 AACACCTGGAAATTTGTCAGAGACTGAAGGGAACTTGTGGAAGCAGTTGA 1379
459 PARAspLeuArqLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
1380 TCCTGATTTGAGAAATAATGTTTAAATCCTAATGCAAAAGATCCCTTG 1429
476 aIValGlyValArqValTrpProGlnAlaIleProGlnPheLeuValI 492
1430 TTTTGGGTGTAAAGATATGCGCAAAAGCCATTCCACAGTCTTGGTGGT 1479
493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
1480 CATTTTGCATCTCCTGTATAGTGCAAAATAATGCGCTCAGAGGATTCG 1529
509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
1530 TCATGGACCTGTTCTTGGGGGCAACTATATGTATGCTGTGGCGCTTGTGG 1579
526 rGcYsValIleGluGlyAlaTyrGluValAlaSerGlyValIleTrpGlyPheLeu 542
1580 GGTGTGTGGAAAGGTCTTACGAGGTTGACGCTGAAGTGAAGAAATTCCT 1629
543 SerArqTyrAlaTyrLys 548
1630 TCACAAATATGCAATACAAA 1647
seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-15
seq_documentation_block:
; Sequence 15, Application US/08808323
; Patent No. 6018105
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Voltrah, Sandra
; APPLICANT: Ward, Eric
; TITLE OR INVENTION: Promoters from Plant
; PROTOPORPHYRINOGEN OXIDASE GENES


```

NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: NO. 6018105artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808.323
FILING DATE:
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38, 241
REFERENCE/DOCKET NUMBER: CGC 1846
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8589
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1826 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum (cotton)
IMMEDIATE SOURCE:
CLONE: pMDC-15 (NRRL B-21594)
FEATURE:
NAME/KEY: misc.feature
LOCATION: 31..1647
OTHER INFORMATION: /product= "Cotton protox-1 coding
OTHER INFORMATION: region"
US-08-808-323-15

alignment_scores:
Quality: 2113.00 Length: 506
Ratio: 4.477 Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249

alignment_block:
US-09-508-418-2 x US-08-808-323-15 ..
Align seg 1/1 to: US-08-808-323-15 from: 1 to: 1826
47 TTPArGThArGcYserValAlalysAspYrThrValProSerSerAl 63
:::|||||
130 TTCAAAGCTCCGAGTCCCTCCGCGAGGGTCCCGATTTCCCTCATCTAA 179
63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
:::|||||
180 AATCGAGGGGAGATCAATCCATCGCGGATTCGCGATTCGATCGTGGAGG 229
80 lYlSerGlyLeuCysIleAlaGluValMetSerAlaAsnTYR..... 94

```

```

|||||.....:
230 GATACAGTGGACTTTCGATTTGCTCAAGCTCTGCCACCAAGCACCGGTGAC 279
95 .....ProAsnLeuMetValIthrGluAlaArgAspAraGlyGlyAs 109
280 GTCGCTTCCATGTGTGTGACGGAGGGCCAGAGACCGTGTGTGGCCAA 329
109 nIleThrThrValGluArgAspGlyTYrLeuTrpGluGluGlyProAsn 126
330 CATCACTACCGTTGAGAGATGATGATCTGTGGAGNAGAGGCCCAACA 379
126 eRPhelInProSerAspPrometLeuThrMetAlaValAspCysGlyLeu 142
380 GTTTCAGCCCTCCGATCCTAATTCFAACCATGGCCGTGATGTGGATT 429
143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
430 AAGGAGATTTGGTTTGTAGTACCCTTAATGACCCGCGATTTGTACTAT 479
159 pLYsGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176
480 GGAAGGAAACTRAGGCTGTGCCCTCCAAAGCAACCGCTTGTCCGTTT 529
176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
530 TTGATTTGATGAGACATTCCTGGAATAAATTAAGCGGTGGGCGCTAAT 579
193 GlyLeuArgProSerProProGlyLysGluGluSerValGluGlnPheVa 209
580 GGATTCGGCCCTCCCTCCGCTGATGAAAGATCGGGGAGAGACTTTGT 629
209 lArGArGAsnLeuGlyGlyValValPheGluArgLeuIleGluProPheC 226
630 GCGCGGTAAATCTTGGTGTAGAGTTTGAACCTTTATGACCAATTTT 679
226 ysSerGlyValIYrValGlyAspProSerLysLeuSerMetLysAlaAla 242
680 GTYCAGGTGTTTATGAGCGAGGATCCTCAAAATTAAGCATGAACACACA 729
243 PheGlyLysValTrpLysLeuGluIthrGlySerIleIleGlyGlu 259
730 TTTGGAGAGATGAGAGCTAGAGAGATTTGGTGCACATCATTGTTGG 779
259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
780 CACTTCAAGACAAATCCAGGAGAGAAATAAAGACACTTAAGCACCCNAG 829
276 sPProArgLeuProLysProLysGlyIthrValGlySerPheArgLys 292
830 ACCCCGCTCGCCAAACCCAAAGGCGCAACGTTGGATCTTTAAGNAG 879
293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
880 GGACTTACCATGCTCTGAGGCAATTTCTAAGAGTTTGGGTGAGCAATGT 929
309 uLYsLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGly 326
930 AAAATTTATCTTGGAAAGCTTTCCAGATTAACCAAAATTTGGCAAGGGGT 979
326 yRHsLeuThrTYrGluThrProGluGlyValValSerLeuGlnSerArg 342
980 AFAACTGTACATTTGAAACACCTGAAAGAAATGATTCCTTCCTCAAGAGA 1029
343 SerIleValMetThrValProSerTYrValAlaSerAsnIleLeuArgPr 359
1030 AGTGTGTAATGACCATTCACCATGTTCCAGTAACCTTGTGGCATCC 1079
359 OleuSerValAlaAlaAlaAspAlaLeuSerAsnPheTYrTrpProPro 376
1080 TCTCTCGCGCTCTGCTGAGATGATTAATCCCAATTTTATATCTCCAG 1129
376 aGLYAlaValThrIleSerTYrProGlnGluAlaIleArgAspGluArg 392
|||||.....:

```

```

1130 TTGCATCAGTCACAGTCTCTCTCCTATCCAAAAGACCCTTCGAAAAGAAATGT 1179
393 LeuValAspIleuLeuLeuLeuLeuLeuLeuLeuHisProArgThrG1 409
1180 TTGATTGATGTGAGTAAAGGGGGTTTGGCCAGTTCACCACCGCAGGCCA 1229
409 ngIyValGluThrLeuGluThrIleuLeuLeuLeuLeuLeuLeuLeu 426
1230 AGGAATTTGAAACTTAAAGGAGATATACAGTTCATCACTTCCCCCAATC 1279
426 rGAlaProLysGlyArgValIleuLeuLeuLeuLeuLeuLeuLeuLeu 442
1280 GAGCTCCATCTGGCAGGGTGTGCTTGTTCCTTAATGCACAAAGATCCTTG 1329
443 AsnProGluIleuLeuSerLysThrGluSerGlnLeuValIleuValAs 459
1330 AACCACTGAAATTTGTCCAAAGACTGCAAGGGGAACTGTAGAACACATG 1379
459 PARAspLeuArgLysMetLeuIleuLeuLeuLeuLeuLeuLeuLeuLeu 476
1380 TCGATATTTGAGAAAATGCTTTAATAATCCTAATGCACAAAGATCCTTG 1429
476 aIValGlyValArgValIleuLeuLeuLeuLeuLeuLeuLeuLeuLeu 492
1430 TTTTGGGTAGAGATAGGGCCAAAGCCATTCACAGTTCTTGTTGGT 1479
493 HisLeuAspThrLeuSerThrAlaLeuAlaLeuMetAsnAspAsnGly 509
1480 CATTTGGCTCTCTTGATAGTGCACAAAATGGCTCCTCAGGGATTCCTGG 1529
509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGly 526
1530 TCATGACACTGTTTCTTGGGGCAACTATGATCTGCTGGCATTAGAC 1579
526 rGcysValGluGlyAlaIleuValAlaLeuSerGlyValIleuLeuLeu 542
1580 GGTGGTGGAGGTCGCTTACGAGGTCGACCTGAAAGTGAAGAAATTCCTG 1629
543 SerArgTyrAlaIleuLys 548
1630 TCACAAATATGATACAAA 1647
seq_name: /cgn2_6/prodata/1/ina/sa_comb.seq:us-09-050-603A-15
seq_documentation_block:
/sequence 15, Application US/09050603A
/GENERAL INFORMATION:
/APPLICANT: Volrath, Sandra
/APPLICANT: Johnson, Marie
/APPLICANT: Potter, Sharon
/APPLICANT: Ward, Eric
/TITLE OF INVENTION: DNA Molecules Encoding Plant
/TITLE OF INVENTION: Protoporphyrinogen Oxidase
/NUMBER OF SEQUENCES: 37
/CORRESPONDENCE ADDRESS:
/ADDRESSER: No. 6023012artis Corporation
/STREET: 3054 Cornwallis Road
/CITY: Research Triangle Park
/STATE: NC
/COUNTRY: USA
/ZIP: 27709
COMPUTER READABLE FORM:
/MEDIUM TYPE: Floppy disk
/COMPUTER: IBM PC compatible
/OPERATING SYSTEM: PC-DOS/MS-DOS
/SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
/APPLICATION NUMBER: US/09/050,603A
/FILING DATE: 30-MAR-1998
/CLASSIFICATION: 800
/PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mels, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1826 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum (cotton)
IMMEDIATE SOURCE:
CLONE: pWDC-15 (NRRL B-21594)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 31..1647
OTHER INFORMATION: /product= "Cotton protox-1 coding
US-09-050-603A-15

alignment_scores:
Quality: 2113.00 Length: 506
Ratio: 4.477 Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249

alignment_block:
US-09-508-418-2 x US-09-050-603A-15 ..
Align seg 1/1 to: US-09-050-603A-15 from: 1 to: 1826
47 TTPARGTHrArGcysSerValAlaLysAspTyrThrValProSerSerAl 63
::::: |||||::::: |||||
130 TTCACGCTCCGATGCTCCCGCCGAGGGTCCACCACTTCCCTCACTAA 179
63 aValAspGlyLyrProAlaIleuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
::::: |||||::::: |||||
180 ATTCGACGGGGGAGAAATCAATCCATCCGCGGATTCGTCATCCTTGGAG 229
80 LyIleSerGlyLeuCysIleAlaGluIleValMetSerAlaAsnTyr..... 94
|||||::::: |||||::::: |||||
230 GATACAGTGGACTTTCATTCCTCAAGCTCTGCGCCACCAAGCACCGTAC 279
95 .....ProAsnLeuMetValThrGluAlaIleuLeuLeuLeuLeuLeu 109
|||||::::: |||||::::: |||||
280 GTCCGCTCCAAATGATGATGACGAGGCGCAGACCGTGGTGGTGCACA 329
109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsn 126
330 CATCACACACCGTTGGAGAGATGATATCTGTGGGAAAGGCGCCACAACA 379
126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
|||||::::: |||||::::: |||||
380 GTTTTCAGCCCTCCGATTCCTAATCTAAACATATGGCCGTGATGAGATTG 429

```

143 LysAspAspLeuValLeuGIyAspProAsnAlaProArgPheValLeuTr 159
 430 AAGCACCATTGGTTTGGTGAAGCCCTAATGCAACCGGATTTGTACTATG 479
 159 PLYSGLYLysLeuArpProValProSerLysLeuThrAspLeuProPheP 176
 480 GGAGGGAAAACCTAAGCCCTGTCCCTCCAAAGCCACCGACTTCCGGT 529
 176 heAspLeuMetSerIleProGlyLysLeuArgAlaIlePheGlyProIle 192
 530 TFGATTTGATGACCATTTGGTGAAGAACTTGGGCTGGGCTTCCGGGCTATT 579
 193 GlyLeuArpProSerProProGlyIleHisGluGluSerValGluGlnPheVa 209
 580 GGGATTTCCGCTCCCTCCCGGCTTAAAGAAATCCGTGGAGGAGTTTGGT 629
 209 IArGATAsnLeuGIyGlyGluValPheGluValArgLeuIleGluProPheC 226
 630 GCCGCCGAAATCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 679
 226 YsSerGIyValIYrValGIyAspProSerLysLeuSerMetLysAlaAla 242
 680 GPTCAGGTGTATTGCAAGGGGATCTTCAAAAATTAAGCATGAAAGCAGCA 729
 243 PheGlyLysValIThrLysLeuGluGluThrGlySerIleIleGlyGlu 259
 730 TTTGGAGAGATATGAAAGAGATGAAAGATGTTGGCAGCATCTTGGTGG 779
 259 YThrPheLysAlaIleLysGluValArgSerSerThrProLysAlaProArgA 276
 780 CACTTTCAAGACAATCCAGAGAGAAATTAAGACACTTAAGCCACCAGAG 829
 276 sPProArGIeuProLysProLysGlyGlnThrValGIySerPheArgLys 292
 830 ACCCGGCTGCCCCAAAACCCAAAGGGCCAAAACAGTTGGATCTTTAGGAA 879
 293 GlyLeuArpMetLeuProAspAlaIleSerAlaArgLeuGIySerLysLe 309
 880 GGCCTTACCAATGCTGCTGAGGCAATGGCAACAGTTGGGATGCAATGT 929
 309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGIyLysGlyIYr 326
 930 AAATATTATCTTGGAGACTTTCCCGATTTTACCAAAATTTGGCCAAATGGAGG 979
 326 YrHisLeuThrYrGluThrProGluGIyValValSerLeuGlnSerArg 342
 980 ATAACTTGACATTTGAAACACCTGAAAGGAAATGGTATCTCTCAGAGATAGA 1029
 343 SerIleValMetThrValProSerYrValAlaIleAspAsnIleLeuArpR 359
 1030 AGTGTGTAAATGACATTCACATCCAGTGGCCAGTAACTGTGTGCATCC 1079
 359 OleuSerValAlaAlaAlaAspAlaLeuSerAsnPheYrYrProProY 376
 1080 FCTCTCGGGCTGCTGCAAGATGCAATTCACCAATTTTATTTATCCCTCCAG 1129
 376 aLGIyAlaValIThrIleSerYrProGluGluAlaIleArgAspGluArg 392
 1130 TTGCATCAGTCAACAGTCTCCATCCAAAAGAAAGCCATTTGAAAAGAAATGT 1179
 393 LeuValAspGIyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGI 409
 1180 TTGATTTGATGGGAGACTTAAAGGGGTTTGGCCAGTGTCCACCCAGCCAGCCA 1229
 409 nGIyValGIuThrLeuGIyThrIleYrSerSerSerLeuPheProAsnA 426
 1230 AGGAATTTGAAACTTTAGGAGCAGTATTAACAGTTCATCAGTTCCTCCCAATC 1279
 426 rglAaProLysGIyArgValLeuLeuLeuAsnTrYrIleGIyAlaLys 442
 1280 GAGCTCCATCTGGCAGGGTGTCTCTTAAACTACCTAAGAGAGAGCTTACC 1329
 443 AsnProGluIleLeuSerLysrThrGluSerGlnLeuValGIyValLys 459

1330 AACACTGGAAATTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379
 459 pArGAspLeuArGIyMetLeuIleLysProLysAlaIleGlnspproLeuY 476
 1380 TCGTGTATTTGAGAAAATGCTTATTAATCCTAATGCAAAAGAAATCCCTTGG 1429
 476 aLValGIyValIArgValTrpProGlnAlaIleProGlnPheLeuValGIy 492
 1430 TTTTGGGTCTAAGACTATAGCCCAAAAGCCATTTCCACACTTCTTGGTGGT 1479
 493 HisLeuAspPThrLeuSerThrAlaLysAlaIleMetAsnAspAsnGlyLe 509
 1480 CATTGGATTCCTCTGATAGTGCAAAATAATGGCTCTCAGAGGATTCGGGTT 1529
 509 uGluGIyLeuPheLeuGIyLysAsnTrYrValSerGIyValAlaLeuGIyA 526
 1530 TCATGAGACTGTTCTTGGGGGCAACTATATGTATCTGGTGTGGCATTTAGGAC 1579
 526 rGcYsValGIuGIyAlaTrpGIyValAlaIleSerGIyValIThrGIyPheLeu 542
 1580 GGTGTGGGAAAGGTGCTTACGAGGTTGGCAGCTGAAAGTGAAGAAATTCCTG 1629
 543 SerArgTrYrAlaTrYrLys 548
 1630 TCACAATATGCAATACAAA 1647
 seq_name: /cgn2_6/protodata/1/ina/6A_COMB.seq:us-09-102-420B-15
 seq_documentation_block:
 ? Sequence ID, Application US/09102420B
 ? Patent No. 6084155
 ? GENERAL INFORMATION:
 ? APPLICANT: Volrath, Sandra
 ? APPLICANT: Johnson, Marie
 ? APPLICANT: Ward, Eric
 ? APPLICANT: Helfetz, Peter
 ? TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
 ? NUMBER OF INVENTION: OXIDASE ("PROFOX")
 ? NUMBER OF SEQUENCES: 43
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: No. 6084155artis Corporation
 ? STREET: 3054 Cornwallis Road
 ? CITY: Research Triangle Park
 ? STATE: NC
 ? COUNTRY: USA
 ? ZIP: 27709
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: PatentIn Release #1.0, Version #1.30
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/09/102,420B
 ? FILING DATE: 22-JUN-1998
 ? CLASSIFICATION: 800
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 09/059,164
 ? FILING DATE: 13-APR-1998
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 09/050,603
 ? FILING DATE: 30-MAR-1998
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 60/126,430
 ? FILING DATE: 11-MAR-1998
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 08/808,931
 ? FILING DATE: 28-FEB-1997
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 60/012,705
 ? FILING DATE: 28-FEB-1996
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 60/013,612

FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Melys, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ. ID NO.: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1826 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
AMTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum (cotton)
IMMEDIATE SOURCE:
CLONE: pMDC-15 (NRRL B-21594)
FEATURE:
NAME/KEY: misc.feature
LOCATION: 31..1647
OTHER INFORMATION: /product= "Cotton protox-1 coding
OTHER INFORMATION: sequence"
US-09-102-420B-15

alignment_scores:
Quality: 2113.00 Length: 506
Ratio: 4.477 Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249

alignment_block:
US-09-508-418-2 x US-09-102-420B-15 ..
Align seg 1/1 to: US-09-102-420B-15 from: 1 to: 1826

```
47 TPTAAGTTRATGCGYSerValAlaLysAspYrThrValProSerSerAl 63
   ::::: |||||:::||||: ||||: |||||
130 TTCAAGCTCCGATGCTCCCTCCGCGAGGGTCCAGGATTTCCCTACTAA 179
   ::|||::: |||||::: |||||
63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
   ::|||::: |||||::: |||||
180 AATCGACGGGAGAAATCATCCATGCCGATTTCCCTACTGGAGAGTGG 229
   ::|||::: |||||::: |||||
80 YLISerGlyLeuCysIleAlaGlnValMetSerAlaAsnYr..... 94
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
230 GATCAGTGACTTTGCATTTGCTCAAGCTCTCCGACCAAGCAAGCCGGTGC 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
95 ..... ProAsnLeuMetValThrcIuAlaLarArgAspArgAlaGlyLys 109
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
280 GTGGTCCCAATGTGATGTGTGACGGAGGCCAAGACCCTGTTGGTGGCAA 329
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109 nIleThrValGluArgAspGlyYrLeuTrpGluGluGlyProAsn 126
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
330 CATCACTACCCTTGGAGAGATGATATCTGTGGGAAADAGGCCCAACA 379
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
126 ePheGlnProSerAspProkMetLeuThrMetAlaValAspCysGlyLeu 142
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
380 GTTTTCAGCCCTCCGATCTATTTAAACCATGGCCGTGATAGAGATTG 429
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
430 AAGGACGATTTGGTTTAGGTGACCCCTAAATGACCCGATTTGTACTATG 476
```

```
159 pLysGlyLysLeuArgProValProSerLysLeuThrasPleuProPheP 176
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
480 GGAGGAAAACCTAAGGCGCTGGCCCTCAAGCCCAAGCCGACTTGGCCCTT 529
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 heAsPleuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
530 TTGATTTGATGACCATTTGCGAAAACCTAAGGCTGGGTTCCGGGCTATT 579
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 GLYLeuArgProSerProproGlyHisGluGluSerValGluGlnPheVa 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
580 GCATTTGCGCCCTCCCTCCGCGGTGTGAAGAAATCGGTGGAGAGATTTGT 629
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 IArgArgAsnLeuGlyGlyValPheGluArgLeuIleGluProPheC 226
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
630 GCGCCGTAATCTTGGTGTGAGGTTTTTGAACGCTTATTTGAACCATTTT 679
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 ysSerGlyValIYrValGlyAspProSerLysLeuSerMetLysAlaAla 242
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
680 GTTCAGGTGTTTATGCAAGGGGATCCCTCAAAAATTAAGCATGAAAAGCA 729
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 PheGlyLysValIrrPlyLeuGluGluIrrGlySerIleIleGlyG 259
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
730 TTTGGAGAGATGAGAGAGCTAGAAAGATTTGGCAGCATCTATTGGTGG 779
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
259 YThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
780 CACTTTCACACAAATCCAGGAGAAATAAGACACCTAAGCCACCCAGAG 829
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
276 sPProArgLeuProLysProLysGlyGlnThrValIleSerPheArgLys 292
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
830 ACCCGCTCCTGCCAAAACCCAAAGGCCAAAACAGTTGATCTTTTAGAAG 879
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 GLYLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerIle 309
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
880 GGACTTACCATTCCTGCTGAGGCAATTCCTACACAGTTTGGGTACAAATGT 929
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyYr 326
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
930 AAAATTTATCTGGAGAGTTCCAGTTATACCAAAAATTTGGCAATGAGAG 979
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
326 YrHisLeuThrYrGluThrProGluGlyValValSerLeuInsSerArg 342
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
980 ATAACTTGACATTTGAAACCTGAAAGGAAATGTTATCTCTGACAGATA 1029
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
343 SerIleValMetThrValProSerYrValAlaSerAsnIleLeuArgPr 359
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1030 AGTGTATTAATGACCATTCATCCCATGTTGCCAAGTAACCTGTTGGATCC 1079
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheYrYrProProV 376
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1080 TCTCTCGGCTGCTGCGACAGATGCATTAATCCCAATTTTATATCCCTCCAG 1129
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
376 aLgLYAlaValIrrIleSerYrProGluGluAlaIleArgAspGluArg 392
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1130 TTGCATAGTACAGCAAGTCTCTATCCAAAAGAACCCATTCGAAAAGAAATGT 1179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
393 LeuValAspGlyLuleuLysGlyPheGlyGlnLeuHisProArgThrG 409
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1180 TTGATTTATGTGTAAGGAGGTTGGCCAGCTTGCACCCAGCGAGCCA 1229
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
409 nGlyValGluThrLeuGlyThrIleYrSerSerLeuPheProAsnA 426
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1230 AGGAATTTGAACCTTTAGGGACGATATACAGTATCATCATTTCGCCAATTC 1279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
426 rGAlaProLysGlyArgValLeuLeuLeuAsnYrIleGlyAlaAlaLys 442
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1280 GAGCTCATCTGCGAGGGTGTCTGTAACATACATAGGAGGAGGTACC 1329
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
443 AspProGluLeuLeuSerLysThrGluSerGlnLeuValGluValAlaLys 459
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1330 AACACGTGGAAATTTGTCACAAAGACTGAAGGAGACTGTTGAAAGCAATTGA 1379
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
```

```

|||||
1380 TCGTCATTTGGAAAAATGCTTATRAAATCTTAATGCAAAAGGATCCTCTTG 1429
476 a1ValGlyValArgValIrrProGlnAlaIleProGlnPheLeuValGly 492
1430 TTTTGGGTFRAAGAGTAAAGGCCAAAGCCATTCACAGTTCCTTGTTGCT 1479
493 HistLeuAspThrLeuSerThrAlaIleValAlaIleMetLysAspAsnGlyLe 509
1480 CATTGGATTCCTTTGATAGTGCACAAAATGCGTCTCAGGGGATTCCTGGGTT 1529
509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
1530 TCANGAGACTGTTCTTGGGGCAACTATGTTCTGGTGGCATTTAGAC 1579
526 rGlyValGluGlyAlaTyrGlyValAlaIleSerGlyValIrrhGlyPheLeu 542
1580 GGTGTGGAGAGGTGCTTACGAGGTTGACGCTGAGTGAAGTGAAGAAATTCCTG 1629
543 SerArgTyrAlaTyrLys 548
1630 TCACAAATTTGCAATACAAA 1647
seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:us-08-472-028A-1

```

```

seq documentation block:
Sequence 1, Application US/08472028A
Patent No. 5767373
GENERAL INFORMATION:
APPLICANT: Ward, Eric R
APPLICANT: Volrath, Sandra
TITLE OF INVENTION: Manipulation of Protoporphyrinogen
TITLE OR INVENTION: Oxidase Enzyme Activity In Eukaryotic Organisms
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,028A
FILING DATE:
CLASSIFICATION: 800
AFORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1748/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1719 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1644
OTHER INFORMATION: /note="Arabidopsis protox-1 cDNA;
OTHER INFORMATION: sequence from pmc-2"
US-08-472-028A-1

```

```

alignment_scores:
Quality: 2077.00      Length: 539
Ratio: 4.230          Gaps: 4
Percent Similarity: 91.095  Percent Identity: 75.139
alignment_block:
US-09-508-418-2 x US-08-472-028A-1
Align seg 1/1 to: US-08-472-028A-1 from: 1 to: 1719
19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 TCCATGAGATTAATCTCTCCGTCACGACACTCAATCCCTCTTCCGTC 77
35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.....ArgT 49
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78 GTTTTCGAAAGCCCAAT..CTCCGATTAATAATGTTATTAAGCCCTTAGAC 124
49 hrArgCysSerValAlaLysAspTyrTrhValProSerSerAlaValAsp 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 TCCGTTGTCAGTGGCCGGTGGACCAACCGCGGATTCCTCAAAAATCCGA 174
66 GlyGlyProAlaAlaGluLeu.....AspCysValIleValGlyAlaG1 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
175 GGGGAGGAGGACCAACCATCACAGCGATTTGTTGTTGTCGGCGGAGG 224
80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
225 TATTTAGTGGTCTTTCATGCTCAGGCGCTTCTCTATAGCAATCCCTGATG 274
95 .....ProAsnLeuMetValIrrhGluAlaArgAspArgAlaGlyLysn 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
275 CTGCTCCGAATTTAATTTGACCCGAGGCTAAGATTCGTTGGAGCAAC 324
110 IleThrTrhValGluIaArgAspGlyTyrIleuTrpGluGluGlyProAsnSe 126
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 ATTTATCAGCTCGTGAAGAGATGTTTCTCTGGAAGAGCCCAAAATAG 374
126 rPheGlnProSerAspPrometLeuThrMetAlaValAspCysGlyLeuL 143
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
375 TTTTCAACCGTGTGATTCCTATGCTCAGTATAGTGTAGATAGTGGTTTGA 424
143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
425 AGGATGATTTGGTGGTGGAGATCCTACGCGCAAGGTTTGTGTGGTGG 474
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
475 AATGGAAATTAAGGCCGGTTCATTCAGAACTAACACAGACTTACCGTCTT 524
176 eaSPLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
525 TGATTTGANTGAGTATTTGGTGGAGAAATTAAGACTGCTTTTGGTGCACCTG 574
193 lYLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
575 GCATTCGACCGTCCACCTCCAGGCTGTAAGAAATCTGTGGAGAGATTTGTA 624
210 ArgArgAsnLeuGlyGlyValAlaPheGluArgLeuIleGluProPheCy 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
625 GGGCGTAACTCGGTGATGAGGTTTTTGAAGCCCTGATTTGAACCGTTTTGG 674
226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaIal 243
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
675 TTCAGGTGTTTATGCTGCTGATCCTTCAAAAATGATGAAACAGCAGCT 724
243 heGlyLysValIrrPylsLeuGluGluTrhGlyGlySerIleIleGlyGly 259
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
725 TTTGGAGAGTTTGGAAACTAGAGCAAAAATGTTGGTGAAGCATATATAGTGGT 774
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProAlaArg 276
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

775 ACTTTTAAGCAATTCAGAGAGAAAAGCCCTCCACAGCAGACAGAGA 824
276 PPROAIVGLeuProLysProLysGlyIuThrValGlySerPheArgLysG 293
825 CCCCGCCCTGCACAAACACAGAGGCCAAACAGTGGTCTTTCAGAGAGG 874
293 IyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
875 GACCTTCAAAAGTTTGCAGAAAGCAATATCTCAGAAATTAAGTAGAAAGTT 924
310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyIuTy 326
925 AAGTGTCTTGAAGAGCTCTCAGATATACRAAAGTGAAGAGCGGAGGATA 974
326 IHisLeuThrTyGluThrProGluIuValSerLeuGlnSerArgS 343
975 CAACCTTAAACATATGAGACTCCAGATGGTTTTCCTCCAGAGCAAAA 1024
343 erIleValMetThrValProSerTrpValAlaSerAsnIleLeuArgPro 359
1025 GTGTTGTAATGACCGTCCATCTCATGTGTCAAGTGGTCTCTCCGCGCT 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTrpTyProProva 376
1075 CTTTCTGAATCTGCTGCAAAAGCACTCTCAAAACTATATATACCCACAG 1124
376 IGIyAlaValThrIleSerTrpProGlnIuAlaIleArgAspGluArgL 393
1125 TGCGACAGATATCTATCTGCTACCCGAAAGAAAGCAATCCGAAACAGAA 1174
393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
1175 TGATATGATGTGAACATAAAGGGTGGGCAATTCATCCAGCAGCAGCAA 1224
410 GIyValGluThrLeuGlyThrIleTrpSerSerSerLeuPheProAsnAr 426
1225 GGAGTGAACATTAAGAACTATCTACAGCTCTCCAGCTTTCACAAATTCG 1274
426 GAlaProLysGlyArgValLeuLeuLeuAsnTrpIleGlyIyAlaLysA 443
1275 CGCACCGCCCGGAAAGAAATTTGCTGTGAACACTACATYGGCGGCTACMA 1324
443 snProGluIleLeuSerSerTrpGlnSerGlnLeuValGluValAlaAsp 459
1325 ACACCGGAATTTCTGTCCAAGTCTGAGAGGTGAGTAGGGAAGCAGTGGAC 1374
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
1375 AGAGATTTGAGAGAAATGCTAATTAAGCCCTAATTCGACCGATCCACTAA 1424
476 IValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
1425 APTAGAGATTTAGGATAGCCCTCAAGCCATTCCTCAGTTCAGTGTGGTCC 1474
493 IHisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
1475 ACTTGAATATCTTGCACACGGCTAAATCATCTCAAGCTCTCCGCGCTAC 1524
510 GIuGlyLeuPheLeuGlyIyAsnTrpValSerGlyValAlaLeuGlyArg 526
1525 GAAGGCGCTAATTTTGGTGGCAATTAACCTGCTGTGTAAGCCCTTAAGCGCG 1574
526 gCysValGluGlyIyAlaTyArgIuValAlaSerGluValThrGlyPheLeuS 543
1575 GTGTGTAGAGAGCGCAATATGAAAACCGCATTTGAGAGCTCAACAACCTTCA 1624
543 erArgTrpAlaLysLys 548
1625 CACGGTAGCGCTTACAAG 1641

```

```

seq_name: /cgn2_6/ptoddata/1/ina/5B_COMB.seq:us-08-808-931-1
seq_documentation_block:

```

```

: Sequence 1, Application US/08808931
: Patent No. 5939602
: GENERAL INFORMATION:
: APPLICANT: Voltath, Sandra
: APPLICANT: Johnson, Marie
: APPLICANT: Potler, Sharon
: APPLICANT: Ward, Eric
: APPLICANT: Helfetz, Peter
: TITLE OF INVENTION: DNA Molecules Encoding Plant
: TITLE OF INVENTION: Protophyryinogen Oxidase and Inhibitor-Resistant Mutants
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5939602artis Corporation
: STREET: 520 white plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591-9005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/808,931
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/012,705
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/013,612
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/020,003
: FILING DATE: 21-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: CGC 1847
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1719 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Arabidopsis thaliana
: IMMEDIATE SOURCE:
: CLONE: pMDC-2 (NRRL B-21238)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 31..1644
: OTHER INFORMATION: /product= "Arabidopsis protox-1"
US-08-808-931-1

```

```

alignment_scores:
Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139

```

```

alignment_block:
US-09-508-418-2 x US-08-808-931-1 ..
Align seg 1/1 to: US-08-808-931-1 from: 1 to: 1719

```


ADDRESS: No. 6018105artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,323
FILING DATE:

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Meligs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1719 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: pMDC-2 (NRRL B-21238)
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1644
OTHER INFORMATION: /product= "Arabidopsis protox-1"
US-08-808-323-1

Alignment scores:
Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139

Alignment_block:
US-09-508-418-2 x US-08-808-323-1 ..
Align seg 1/1 to: US-08-808-323-1 from: 1 to: 1719

19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSer 35
28 TCACAGGAGGTATCTCTCTCCGCGACGACTCAATGCGCTTCCGCTC 77
35 rIleSerLysArgAsnSerValAsnCyAsnGlyTrp.....ArgT 49
78 GTTTTCGAAGCCCAAF...CTCCGATTAATGTATATAAGCCCTTTAGAC 124
49 hrArgCysSerValAlaLysAspTrpThrValProSerSerAlaValasp 65
125 TCCGTTGTTCAGTGGCCGCGTGAGCCACCACCGTCGATCTTCAAAAATCGAA 174

66 GlyGlyProAlaAlaGluLeu.....AspCysValIleValGlyAlaG 80
175 GCGGGAGAGGACACACATACGACGAGATTGTGTGATTTGCGGGGAGG 224
80 YIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
225 TATATAGTGTCTTTCATCCGCTCAGGCCCTTGCTACATCAAGCATCC 274
95ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyIAsn 109
275 CTGCTCCGAAATTAATTTGTGACCGGAGGCTAAGGATCGTGGAGCC 324
110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsn 126
325 ATTATCACTCGTAGAGAGAATCGTTTCTCTGGGAGAAAGGCCCAATAG 374
126 rPheGlnProSerAspPrometLeuThrMetAlaValAspCysGlyLeu 143
375 TTTTCAAACCGCTGATCCATGCTCACAATGATGATGATGATGATGAT 424
143 yAspAspLeuValIleuGlyAspProAsnAlaProArgPheValIleu 159
425 AGGATGATTTGGTGTGGAGATCCCTACTGCGCCAAAGCTTTGTGTGG 474
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176
475 AATGGGAANTGAGGCGGCTTCATCGAAGCTAACACAGACTTCCGTTCT 524
176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 193
525 TGATTTGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 574
193 LysLeuArgProSerProGlyHisGluGluSerValGluGluIlePhe 209
575 GCATTCGACCGCTCACCTCCAGGCTCGAAGAAATCGTGGAGGAGTTG 624
210 ArgArgAsnLeuGlyGluValPheGluArgLeuIleGluProPheGy 226
625 CGCGCTAACCTCGTGTGATGAGGTTTGGAGCCCTGATTTGAACCGT 674
226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 243
675 TTCAGGTGTTTATGCTGTGATCTTCAAACACTGACATGAACAGCAG 724
243 heGlyLysValTrpLysLeuGluGluThrGlyLysIleIleGlyGly 259
725 TTGGGAAGGTTTGGAAATGAGCAAAATGTCGAGGCAATTAATAGTGT 774
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg 276
775 ACTTTTAAGCAATTCAGGAGGAAACCGCTCCAAAGGCAAGACGAGA 824
276 pProArgLeuProLysProLysGlyInThrValGlySerPheArgLysG 293
825 CCGCGCCCTCCAAACACACAGGCGCAACAGTGGTCTTTCAGGAGAG 874
293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLys 309
875 GACTTCGATGTTGCCAGAGCAATATCTCAGAAATTAAGTATGCAAAAG 924
310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyTyr 326
925 AAGTTTGTCTTGGAAAGCTCTCAGGTATCACAACCTGAGCGGAGGAGA 974
326 rHisLeuThrTyrGluThrProGluGlyValIleSerLeuGlnSerArg 343
975 CAACCTTAACATATGAGACTCCAGATGTTAGTTCCTCCGTCACAGGAAA 1024
343 rIleValIleMetThrValProSerTyrValAlaSerAsnIleLeuArg 359
1025 GTGTTGTAATGAGCGGTCCAMTCAATGTTGCAAGTGGTCTTCCGCCC 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProPro 376


```

1075 CTTTGAATGCTGCAATACCTCAAAACATATTTTACCACCAGT 1124
1076 IGLYALVAlThrIleSerTYrProGIInGluAlaIleArgAspGIuArgL 393
1125 TCGAGAGATATCTATCTGTACCCGAAGAACAATCCGAACAGAAATGTT 1174
393 euValAspGIuLeuLysGIyPheGIyGInLeuHisProArgThrGIn 409
1175 TGTATGATGTGAACTAAAGGGTTTTGGGCAATGTCATCCAGCCAGCAA 1224
410 GLYValGIuThrLeuGIyThrIleTYrSerSerIleuPheProAsnAr 426
1225 GGAAGTTGAAACTTGTGAACTATGACAGCTCCACTCTTCCAAACG 1274
426 GAlaProLysGIyArgValIleuLeuLeuAsnTYrIleGIyAlaLysA 443
1275 CCGACCCGCCGGAAGAAATTTGGCTGTGAACACTTGGCGGGCTTACAA 1324
443 snProGIuIleLeuSerTYrThrGIuSerGIuLeuValGIuValAsp 459
1325 ACACCGGAATTCGTCCAAAGCTGTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374
460 ArgAspLeuArgLYsMetLeuIleLysProLYsAlaGIuAspProLeuVa 476
1375 AGAGATTTGAGGAAATACTTAATTAAGCTTAATTCGACCGAATCACTTAA 1424
476 IValGIyValArgValItrProGIInAlaIleProGIInPheLeuValGIyH 493
1425 ATTAGAGATGTAGGTATGGCCCTCAAGCCATTCACAGTTTCTTACTGTGTC 1474
493 ILeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGIyLeu 509
1475 ACTTTGATATCTTGCACACCGCTAAATCCTCAACGCTCTTCCGGCTTAC 1524
510 GIuGIyLeuPheLeuGIyLYsAsnTYrValSerGIyValAlaLeuGIyAr 526
1525 GAAGGGCTATTTTGGTGGCAATTTAGCTCCCTGTGTAGCCCTTAGGCGC 1574
526 gCysValGIuGIyAlaTYrGIuValAlaSerGIuValItrGIyPheLeu 543
1575 GTGTGTGAAAGCGCATATGAAACCGCGAATGTGAGGTCACAAACTTCAATGT 1624
543 eArgTYrAlaTYrLys 548
1625 CACGGTACGCTTACAG 1641
seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:us-09-050-603A-1

```

```

seq_documentation block:
; Sequence 1. Application US/09050603A
; Patent No. 6023012
; GENERAL INFORMATION:
; APPLICANT: Volirath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6023012artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050, 603A
FILING DATE: 30-MAR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808, 931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012, 705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013, 612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020, 003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38, 241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1719 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: pMDC-2 (NRRL B-21238)
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1644
OTHER INFORMATION: /product= "Arabidopsis protoc-1"
US-09-050-603A-1

```

```

alignment_scores:
Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139
alignment_block:
US-09-508-418-2 x US-09-050-603A-1 ..

```

```

Align seg 1/1 to: US-09-050-603A-1 from: 1 to: 1719
19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 TCCATGAGATTAATCTTCTCCGTCGACGACTCAATGCTTCTTCCGTC 77
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
35 rIleSerLysArgAsnSerValAsnCysAsnGIyTrp.....ArgT 49
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
78 GTTTTCGAAGCCCAAT...CTCCGATTAATGTATTAAAGCCTCTTAGAC 124
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
49 hrArgCysSerValAlaLysAspTYrThrValProSerSerAlaValAsp 65
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 TCCGTTGTTCAGTGGCGGTGACCAACCGCGGATCTTCAAAAATCGAA 174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 GIYGIYProAlaIaGIuLeu.....AspCysValIleValGIyAlaGI 80
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
175 GCGGAGGAGGACCAACCACTACGACGGATGTGTGATTTGGGGGAGG 224
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 rIleSerGIyLeuCysIleAlaGIuValMetSerAlaAsnTYr..... 94
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
225 TATGTGTGTCTTTGCAATCGCTCAGCGCGCTTGCTACTAGACATCTGATG 274
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

95 ..... ProasnlleuMetValThrGluLaaArgAspArgAlaGlyGlyAsn 109
275 CTGCTCCGAATTTAATGTGTACCCAGAGGCTAAGGATCGTTGGGGGCAAC 324
110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluIleProAsn 126
325 ATTTATCTGCTGTGAAGAAATGGTTTCTGGGAGAGAGTCCCAATAG 374
126 rPheGlnProSerAspPromeLeuThrMetAlaValAspCysGlyLeuL 143
375 TTTTCAACCCGCTGATCTATGCTCACTATGCGGTGATAGTGTGTTGA 424
143 ysaAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
425 AGGATGATTTGGTGTGGAGATCTCTACTGCCCAAGTTTGTGTGG 474
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176
475 AATGGGAATTTGAGGCCGGTCCATCCAGCTAAGACACTTACCGTCTT 524
176 eaSPLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
525 TGATTTGATGAGTATTTGGGAAAGATTAAGAGCTGTTTGTGTCACCTTG 574
193 LysLeuArgProSerProGlyLysGlnGluIleGluIleGluGlnPheVal 209
575 GCATTCGACCCGTCACCTCCAGGTCGTGAAGATCTGTGGAGAGATTGTA 624
210 ArgArgAsnLeuGlyGlyValPheGlnLysLeuIleGluProPheCys 226
625 CGGGCTAACCCTCGGATGAGGATGTTTGTAGCCCGCTGATTTGAACCGT 674
226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
675 TTCAGGTGTATCTGTGTCCTGATCCCTTCAAAACTGAGCATGAAAGGAG 724
243 heGlyLysValTrpLysLeuGlnGluIleThrGlySerIleGlyGly 259
725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGCTGGAAAGCATTAATAGTGT 774
260 ThrPheLysAlaIleLysGlnLysSerSerThrProLysAlaProArgAs 276
775 ACTTTTAAAGGCAATTCAGAGAGGAAAAACGCTCCCAAGCAGAAAGCAGA 824
276 pProArgLeuProLysProLysGlyIleThrValGlySerPheArgLysG 293
825 CCCGGCCCTCCAAAACCAAGGCCAAAACAGTTGGTCTTTCAGGAAAGG 874
293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
875 GACTTCGAAATGTTGCCAAGCAAAATCTGCCAAGATTTAGTGTGCAAAAGT 924
310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGlyLysGlyTyr 326
925 AAGTGTCTTGGAAAGCTCTCAGGATCACTAAGCTGAGAGCGGAGAGATA 974
326 rHisLeuTrpTyrGluThrProGlnGlyValValSerLeuGlnIleSerArg 343
975 CAACSTTAAACATGATGAGACTCCAGATGGTTTAACTTCCCGGAGCAGCAADA 1024
343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
1025 GTGTGTAAATGACGGGTGCATGTCAMTGGTGCAAAGTGTCTTGGCCCT 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
1075 CTTTCGAAATCTGTCGCAAAATGACACTCCAAAATAATTAATCCSCCAAGT 1124
376 IeGlyAlaValAlaThrIleSerTyrProGlnGluLysIleLeuArgAsp 393
1125 TCCAGGAGATATCTATCTGTTACCCGAAAGCAAAATGCAACAGAAATGTT 1174
393 euValAspArgLysLeuLysGlyPheGlnLeuHisProArgThrGln 409

```

```

1175 TGATGATGTGTAACTAAAGGTTTGGGCAATTCGATCCAGCGCAAGAA 1224
410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
1225 GAGGTTGAAACATTTAGGAACATTAACAGCTTCACTCTTCCAAAATCG 426
426 gaIaProLysGlyArgValLeuLeuLeuAsnTrpIleGlyGlyValAlaLysA 443
1275 CGCACCGCCGGAAAGATTTTGTGTGAACTAATTTGGCGGGTCTACAA 1324
443 snProGluIleLeuSerLysThrGlnSerGlnLeuValGluValAlaAsp 459
1325 ACACCGGAATTTCTGTCCAAAGTCCGAAAGTGTAGTGTGGAACAGTGTGC 1374
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
1375 AGAGATTTGAGGAAATGCTAATTAAGCCTAATTCGACCGACTCCATTA 1424
476 ValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGln 493
1425 AITTAGGAGTTTAGGTTATGCTTCMAAGCCATTCCTCAGTTTCTTAGTGTGC 1474
493 IsLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
1475 ACTTTGATATCTTGACACCGGCTAAATTCATCTTAACGCTTCGGGCTAC 1524
510 GlnGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
1525 GAAAGGCTATTTTGGTGGCAATTAAGTCAGCTGCTGTAGCCCTTAGAGCCG 1574
526 gCysValGluGlyAlaTyrGlnValAlaAserGluValIleThrGlyPheLeu 543
1575 GTGTGTAGAAAGGCGCAATTAAGAAACCGGATTAAGGTCACAACTTCATGT 1624
543 erArgTyrAlaTyrLys 548
1625 CACGGTACGCTTACAAAG 1641
seq_name: /gn2_6/ptodata/L/lna/6A_COMB.seq:US-09-102-4208-1
seq_documentation_block:
; Sequence 1, Application US/091024208
; Patent No. 6084155
; GENERAL INFORMATION:
; APPLICANT: Volirath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Ward, Eric
; APPLICANT: Helicz, Peter
; TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 6084155artis Corporation
; STREET: 3054 Cornwails Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,4208
; FILING DATE: 22-JUN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/059,164
; FILING DATE: 13-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,603

```



```

1125 TGCACGAGTACTACTCTGTRACCCGAAAGAAAGCAATCCGAAAGAAATGTT 1174
393 euValaapglYgluLeuLySgLyphEgLygluLeuHISProaHgrThGln 409
1175 TGAATAGATGGTGAACCTAAGGGGTTTGGCAATTCACCGACCGACGAA 1224
410 GllYValgluThrlEugLyThrIleTyrSerSerLeupheProaSer 426
1225 GGAAGTTGAATATTAGGACTATCTTACAGCTCCCTCCTCTTTCCAAATCG 1274
426 gAlaProLySgLyArYValLeuLeuLeuAAsnTYrIleGlyglYAlaLysA 443
1275 CGCACCGCCCGGAGAAATTTTGGCTGTAACATTCAGTTGGCGGGCTFCAAA 1324
443 snProGluIleLeuSerLySthrGluSerGlnLeuValGluValAlaASP 459
1325 ACACCCGAAATTCGTCCAAAGCTGAAGGTGAGTTAGTGAAGGAGTTGAC 1374
460 ArgAspLeuArGlyMetLeuIleLySProLySAlaGlnAspProLeuVa 476
1375 AGAGATTGGAGAAAGCTAATTAATTAAGCCTTAATTCGACCCATCCACTAA 1424
476 lValGlyValArYValTrpProGlnAlaIleProGlnPhELeuValGlyH 493
1425 AFTAGGAGTFTAGGGATGGCTCAAGCCATTCCTCAGTTCCTAGTGTGATC 1474
493 lSLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
1475 ACTTGGATATCTGTGACACGGCTAAMTCATCTCAACGCTCTCGGGCTAC 1524
510 GluGlyLeuPheLeuGlyGlyAsnTYrValSerGlyValAlaLeuGlyAr 526
1525 GAAGGGCTATTTTTGGTGGCAATTCAGCTCGCTGTGTAGCCCTTAGAGCGC 1574
526 gCyValGluGlyAlaTYrGluValAlaSerGluValThrGlyPheLeuS 543
1575 GTGTGTAGAAAGCCGATATGAACCGCGATTTAGGCTCAACAACTTCATGT 1624
543 eArGTYrAlaTYrLys 548
1625 CACGGTACGCTTACAAAG 1641
seq_name: /cgn2_6/prodata/1/ina/68_COMB.seq:US-09-071-296-1

```

```

seq documentation block:
? Sequence 1, Application US/09071296
? Patent No. 6177245
? GENERAL INFORMATION:
? APPLICANT: Ward, Eric R
? APPLICANT: Voltrath, Sandra
? TITLE OF INVENTION: Manipulation of Protoporphyrogen
? TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms
? NUMBER OF SEQUENCES: 12
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Ciba-Geigy Corporation
? STREET: 7 Skyline Drive
? CITY: Hawthorne
? STATE: NY
? COUNTRY: USA
? ZIP: 10532
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/071.296
? FILING DATE: 06-JUN-1995
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/261,198
? FILING DATE: 16-JUN-94
? ATTORNEY/AGENT INFORMATION:

```

```

? NAME: Elmer, James Scott
? REGISTRATION NUMBER: 36,129
? REFERENCE/DOCKET NUMBER: CGC 1748/CIP
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 919-541-8614
? TELEFAX: 919-541-8689
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1719 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHEICAL: NO
? ANTI-SENSE: NO
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 31..1644 /note="Arabidopsis protox-1 cDNA,
? OTHER INFORMATION:
? OTHER INFORMATION: sequence from PMDC-2"
?
? US-09-071-296-1

```

```

alignment_scores:
Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139

```

```

alignment_block:
US-09-508-418-2 x US-09-071-296-1
Align seq 1/1 to: US-09-071-296-1 from: 1 to: 1719

```

```

19 SerSerProLeuAlaPheLeuAsnArGTrhSerPheIleProPheSerSe 35
28 TCcATGGAGTATCTCTCTCCGCGAGCACTCAATCCCTCTCCGTC 77
35 rIleSerLySArgAsnSerValAsnCySAsnGlyTrp.....ArgT 49
78 GTTTGGAAAGCCCAAT...CTCCGATTAATGTTTAAAGCCCTTAGAC 124
49 hrArgCySerValAlaLysAspTYrThValProSerSerAlaValaASP 65
125 TCCGTTGTTCCAGTGGCCGGTGGACCAACCGGATCTTCAAAAATCGAA 174
66 GLYGlyProAlaAlaGluLeu.....AspCySValIleValGlyAlaG1 80
175 GCGGGAGAGGACACCAACACAGCAGCGANTGTGTGATTTGGCGGAGG 224
80 yIleSerGlyLeuCySIIeAlaGlnValMetSerAlaAsnTYr..... 94
225 TATTAGTGGCTTGTGCTCAGCGGCTTCCCTACTAGCAATCCGATG 274
95 ....ProAsnLeuMetValThrGluAlaArGAspPrGAlaGlyAsn 109
275 CTGCTCCGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 324
110 IleThrThrValGluArGAspGlyTYrLeuTrpGluGluGlyProAsnSe 126
325 ATTATACCTGCTGTAAGAAATGTTTCTCTGGGAAAGAGGCTCCAAAT 374
126 rPheGlnProSerAspPrometLeuThrMetAlaValaASPcySgLyLeu 143
375 TTTTCAACCGCTGATCCCTATGCTCACTAAGGAGTGAAGTGGTTGA 424
143 ySAspAspLeuValIleGlyAspProAsnAlaProArGpPheValLeuTrp 159
425 AGGATGATTTGGTGGGAGATCCACTACGCCCAAGGTTTGGTGTGG 474
160 LySgLyLysLeuArGProValProSerLySLeuThrAspLeuProPhePh 176
475 AATGGAAATTTGAGGCGGTTCCATCGAAGCTAAACAGACTTACCGTCTT 524

```

176 aaSPLeuKtSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
 525 TGATTTGATGAGTATGTTGGTGGAGATTAGAGCTGGTGGTGGTCCACTTG 574
 193 LLeuArgProSerProProlGlyHisGluGluSerValGluGlnPheVal 209
 575 GCATTCGACCCCTCACCTCCAGGTCGTGTAAGAAATGTGGAGAGTGTGTA 624
 210 ArGArGAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
 625 CGGGGTAACCTCGGTGATGAGGTTTGGAGCCCTGATTTGACCGCTTTG 674
 226 sSerGlyValIyrValGlyAspProSerLysLeuSerMetLysAlaLap 243
 675 TTCAGGGTGTATGTCTGGATCCCTTCAAAAAGTGAAGTGAAMAACACCG 724
 243 heGlyLysValIyrPlyLysLeuGluGlnIyrHisGlySerIleIleGly 259
 725 TTGGGAAGGTTTGGAAACTAGACGAAATGGTGGAAACATATATATGGTGT 774
 260 ThrPheLysAlaLalIeLysGluArgSerSerThrProLysAlaProArg 276
 775 ACTTTTAAAGCAATTCAGAGAGGAAACACGCTCCCAAGCAGACAGAGA 824
 276 pProArgLeuProLysProLysGlyGlnIyrValGlySerPheArgLysG 293
 825 CCGCGCGCTGCAAAACCAACGAGGCAAAACAGGTTGTTTCCAGGAAAG 874
 293 LLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
 875 GACTTCAAGTGTCCAGAAAGCAATATCTGCAAGATTAGGTAGCAAAAGT 924
 310 LysLeuSerIyrPlyLysLeuSerIleThrLysSerGluLysGlyGly 326
 925 AAGTGTGCTGGAGAGCTCAAGGATCACTAAGGCTGAGGAGCGAGAGATA 974
 326 rHisLeuThrIyrGluIyrProGluGlyValValSerLeuIleSerArg 343
 975 CAATTTAAATATAGAGAGACTCCAGATGGTTTTCCTGACAGGCAAAA 1024
 343 erIleValMetThrValProSerIyrValAlaSerAsnIleLeuArgPro 359
 1025 GTGTGTAAATGACGGTCCCATCTCATGTGTCAGTGTCTCTTGGCGGCT 1074
 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheIyrIyrProProVa 376
 1075 CTTTCTGAATCTGCTGCAAAATGCACTCTCAAAAACATATATATACCCAC 1124
 376 lGlyAlaValIyrIleSerIyrProGluGlnIalaIleArgAspGluArgL 393
 1125 TGACGACGATTCATCTCTGATCCCGAAAGCAAAATCCGACAGAAATGTT 1174
 393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgIyrGln 409
 1175 TGATATAGTGGTGAACATPAAAGGGTGGGCAATTTGGCATCCACGACG 1224
 410 GlYValGlnIyrLeuGlyIyrIleIyrSerSerSerLeuPheProAsnAr 426
 1225 GAGGTGAAACATTAGAACTATCTACAGCTCTCACTCTTCCAAATATG 1274
 426 gAlaProLysGlyArgValLeuLeuLeuAsnIyrIleGlyValAlaLysA 443
 1275 CGACCGCCCGGAAAGAAATTTGGTGTGTAACATACATTTGGGGGTGTACA 1324
 443 snProGluIleLeuSerLysIyrHisGlnSerGlnLeuValGluValAla 459
 1325 ACACCGGAATTTCTCCAAAGTCCGAAAGGTGAGTGTGGAGACAGTTGAC 1374
 460 ArGAspLeuArGlyMetLeuIleLysProLysAlaGlnAspProLeuVa 476
 1375 AGAGATTTGGGAAATGCTAAATTAAGCCCTAATTCACACCGATCCACTTAA 1424
 476 lValGlyValArgValIyrProGlnAlaIleProGlnPheLeuValGlyH 493

```

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:us-08-808-931-11
seq documentation block:
? Sequence 11 Application US/08808931
? Patent No. 5939602
? GENERAL INFORMATION:
? APPLICANT: Volrath, Sandra
? APPLICANT: Johnson, Marie
? APPLICANT: Potter, Sharon
? APPLICANT: Ward, Eric
? APPLICANT: Helletz, Peter
? TITLE OF INVENTION: DNA Molecules Encoding Plant
? TITLE OF INVENTION: Protooporphyrinogen Oxidase and Inhibitor-Resistant Mutants
? NUMBER OF SEQUENCES: 35
? CORRESPONDENCE ADDRESS:
? ADDRESS: No. 5939602arlls Corporation
? STREET: 520 White Plains Road, P.O. Box 2005
? CITY: Tarrytown
? STATE: NY
? COUNTRY: USA
? ZIP: 10591-9005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/808,931
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/012,705
? FILING DATE: 28-FEB-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/013,612
? FILING DATE: 28-FEB-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/020,003
? FILING DATE: 21-JUN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Meigs, J. Timothy
? REGISTRATION NUMBER: 38,241
? REFERENCE/DOCKET NUMBER: CGC 1847
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (919) 541-8587
? TELEFAX: (919) 541-8689
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1847 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA

```

? HYPOTHETICAL: NO
 ? ORIGINAL SOURCE:
 ? ORGANISM: soybean
 ? IMMEDIATE SOURCE:
 ? CLONE: PMDC-12 (NRRL B-21516)
 ? FEATURE:
 ? NAME/KEY: CDS
 ? LOCATION: 55..1683
 ? OTHER INFORMATION: /product="soybean protox-1"
 ? US-08-808-931-11

alignment_scores: Quality: 2062.50 Length: 550
 Ratio: 4.218 Gaps: 7
 Percent Similarity: 88.909 Percent Identity: 73.273

alignment block:
 US-09-508-418-2 x US-08-808-931-11 ..
 Align seg 1/1 to: US-08-808-931-11 from: 1 to: 1847

```

3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
   :::::|||||:  |||||:::| | | | | | | | | | | | | | | | |
102 TCTTGGCCCTCCCTCCATTCGCCAACCTCTTCTTCCACCTCCCTCC 151
   :::::|||||:  |||||:::| | | | | | | | | | | | | | | | |
19 rSerProLeuAlaPheLeuAsnArGThrSerPheIleProPheSerSe 36
   :::::|||||:  |||||:::| | | | | | | | | | | | | | | | |
152 GAAA.....ATTCCC.....ATTCCC.....T 162
36 LeSerLysArgAsnSerValAsnCysAsnGlyTrpArGThrArgCysSe 52
   |||||:|||||:  |||||:|||||:  |||||:|||||:  |||||:|||||:
163 CGCTCTGCCCTAACCTATT.....CTACGCTCTCTC 194
52 rValAlaLysAspTrpThrVal...ProSerSerAlaValAspGlyLyp 68
   :::::|||||:  |||||:| | | | | | | | | | | | | | | | |
195 CATTCGGGAGGAAATCCACCGCGCTCCGCCCAAAACAGAGACTCCGCC 244
68 rOAlaAlaGluLeuAspCysValIleValAlGlyAlaGlyIleSerGlyLeu 84
   :::::|||||:  |||||:| | | | | | | | | | | | | | | | |
245 CC.....GTGGACTGCTGCTGCTGGGGGAGGCGGTACAGGCGCTC 285
85 CysIleAlaGlnValMetSerAlaAsnTrp.....ProAsnLeuMetVal 99
   |||||:|||||:  |||||:|||||:  |||||:|||||:  |||||:|||||:
286 TGCATTCGCCCAAGCCCTCCACCAACACGCAATGCCAATGCCATGCTG 335
99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValAlaArg 116
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
336 CACGGAGGCGCCGAGNCCGCTGCGCGGCACATCACCACGATGGAGAGG 385
116 spGlyTrpLeuTrpGluGlnGlyProAsnSerPheGlnProSerAspPro 132
   |||||:|||||:  |||||:|||||:  |||||:|||||:  |||||:|||||:
386 ACGGATACCTCTGGGAAAGAGGCCCCACACACACTTCCACCTCTTGATCA 435
133 MetLeuTrpMetAlaValAspCysGlyLeuLysAspAspLeuValLeuG 149
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
436 ATGCTCACCAATGATGGTGGACAGAGGTTAAAGGATGAGCTTGTGGGG 485
149 yAspProAsnAlaProArGpHeValLeuTrpLysGlyLysLeuArGpV 166
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
486 GGATTCATGATGACCTCCGTTGTGTGGACAGCAAGATTTGAGGCGCG 535
166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
   |||||:|||||:  |||||:|||||:  |||||:|||||:  |||||:|||||:
536 TGCCCGGAGAGCTGACGTGATTTGCCCTTCTTTCCTTGACTTGATGACAT 585
183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArGpProSerPro 199
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
586 GGCAAAATCAGGGGCTGGCTTGGTGGCTGGAAATTCGGCCTCCCTCC 635
199 oGlyHisGluLeuSerValGluGlnPheValAlaArgAsnLeuGlyG 216
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
636 AGTCATATGAGAAATCGTGTGAAGATTTGTCGTCCGGAACCTTGGTATG 685
  
```

```

216 lValPheGluArgLeuIleGluProPheCysSerGlyValTrpTrValGly 232
   |||||:|||||:  |||||:|||||:  |||||:|||||:  |||||:|||||:
686 AGGTTTGAACGGTGTAGAGCCCTTTTGTTCAGGGGTATGTCAGGGC 735
223 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
736 GATCCTTCAAAATTAAGTATGAAACACAGCATTCGGGAAAGTTTGGAACT 785
249 uGluGlnThrGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
786 GAAAAAAATGGTGGATGATTTGGTGGAACTTCAAGCAAAATACAAAG 835
266 lArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
   |||||:|||||:  |||||:|||||:  |||||:|||||:  |||||:|||||:
836 AAGAAATGGAGCTTCMAAAACCAACCTCCGAGATCCCGCTTCGCCAAAACA 885
283 LysGlyGlnThrValGlySerPheArgGlyLysGlyLeuArgMetLeuProAs 299
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
886 AAAGTCCAGACCTGTGGATCTTCCGGAAAGGACTTACCAGTGTCCCTGCA 935
299 palAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeu 316
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
936 TGCATTTCTGCCAGACATGAGCAACAAGTAAAGTATCTTGGAAAGCTTT 985
316 eSerIleThrLysSerGluLysGlyLysGlyTrpHisIleuThrTrpGluThr 332
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
986 CAAGTATTTGTAACCTGATGATGATGAGAGAGTACAGTTTGACATATGAAACA 1035
333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1036 CCAGAGAGAGTGGTGTCTTTCGACCAAAACCTGTTCCTGACCACTCC 1085
349 oSerTrpValAlaIleSerAsnIleLeuArgProLeuSerValAlaAlaAla 366
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1086 TTCCTATGTTGCTAGTACATGCTGGCTCTGCTGCTGCTGCTGCTGCTG 1135
366 spAlaLeuSerAsnPheTrpTrpProValAlaValAlaThrIleSer 382
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1136 ATGCATTTCAAAAGTTTATTTACCCCTCCAGTGGCTGCAAGTCCAAATACC 1185
383 TrpProGluGlnAlaIleArgAspGluArgLeuValAspGlyLysLeu 399
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1186 TATCCAAAAGAGCTTATGATCAGAAATGCTGATGATGATGATGATGATG 1235
399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValAlaGluThrLeuGly 416
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1236 GGGGTTTGGTCAATTCATCCACGCTAGCCCAAGAGGTAACATTAAGGA 1285
416 hTrIleTrpSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1286 CTATATATACAGCTCATCATTCCTCCCAACCGAGCACCTCGGAAGGGT 1335
433 LeuLeuLeuAsnTrpIleGlyValAlaLysAsnProGluIleLeuSerLys 449
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1336 CTACTCTTGAATTAACATTTGGAGGAGCAACTAATCTGGAATTTTATCGGA 1385
449 sThrGluSerGlnLeuValGluValAlaAspArgAspLeuArgLysMetL 466
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1386 GACGGACAGTGAACCTGTGGAACAACATGATCGAGATTTGGAGAAATACC 1435
466 euIleLysProLysAlaGlnAspProLeuValAlaGlyValArgValTrp 482
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1436 TTATTAACCCAAATGCCCAAGATCCATTTGTGTGGTGGAGTCCGAGTCC 1485
483 ProGluAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1486 CCTCAAGCTATTTCCACAGTTCATGTTGGCATTTGATCTTCATGATGAT 1535
499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
   |||||:|||||:  |||||:| | | | | | | | | | | | | | | | |
1536 TGCTAAAGCTTCTATGCAAAATACCTGGGTTTGAAGGGGCTTCCCTGGGG 1585
  
```

516 lYAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
 |||||||
 1586 GAAATTAATGTTCTGGTGTTCCTGGGACGATCGGTTGAGGAGCCCTAT 1635
 533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
 |||||||
 1636 GAGTAGACGCTGAAGTAAAGCATTTTCTCACAATATAGAGTACAAA 1683

seq_name: /cgn2_6/ptodata/1/ina/fa_COMB.seq:US-08-808-323-11

seq_documentation_block:

; Sequence 11, Application US/08808323
 ; Patent No. 6018105
 ; GENERAL INFORMATION:
 ; APPLICANT: Johnson, Marie
 ; APPLICANT: Volrath, Sandra
 ; APPLICANT: Ward, Eric
 ; TITLE OF INVENTION: Promoters from Plant
 ; TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 6018105artis Corporation
 ; STREET: 520 White Plains Road, P.O. Box 2005
 ; CITY: Fairylowm
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10591-9005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/808, 323
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/012, 705
 ; FILING DATE: 28-FEB-1996
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/013, 612
 ; FILING DATE: 28-FEB-1996
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/020, 003
 ; FILING DATE: 21-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meigs, J. Timothy
 ; REGISTRATION NUMBER: 38, 241
 ; REFERENCE/DOCKET NUMBER: CGC 1846
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (919) 541-8587
 ; TELEFAX: (919) 541-8689
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1847 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: soybean
 ; IMMEDIATE SOURCE:
 ; CLONE: pMDC-12 (NRRL B-21516)
 ; NAME/KEY: CDS
 ; LOCATION: 55..1683
 ; OTHER INFORMATION: /product= "soybean protox-1"
 ; US-08-808-323-11

alignment_scores: Quality: 2062.50 Length: 550

Ratio: 4.218 Gaps: 7
 Percent Similarity: 88.909 Percent Identity: 73.273
 alignment_block:
 US-09-508-418-2 x US-08-808-323-11
 Align seg 1/1 to: US-08-808-323-11 from: 1 to: 1847

3 ThrThProIleAlaAsnHsProAsnIlePheThHsGlnSerSerse 19
 ..|||
 102 TCTTGGCCCTCCCTCCATTCCTCCCAACCTTTCTTCCCTCCCTCCACCC 151
 19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSeri 36
 |||
 152 GAA.....ATGCC.....T 162
 36 le.SerLysArgAsnSerValAsnCysAsnGlyTTPArgThrArgCysse 52
 |||
 163 CGCTCTCGCCCTTAACCTTAT.....CTACGCTGCTC 194
 52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
 |||
 195 CATTTGGGAGGAAATCCACCGCGTCTCCGCCAAACCAAGAGACTCCGCC 244
 68 rAlaAlaGluLeuAspCysValIleValIleGlyAlaGlyIleSerGlyLeu 84
 |||
 245 CC.....GTGACCTGCTGCTGCTGGGAGGCGCTACGCCCTC 285
 85 CysIleAlaGlnValMetSerAlaAsnTyr.....ProAsnLeuMetVa 99
 |||
 286 TGCATCGCCACAGCCCTCCGCCAACAACGCCAATGCCAATGCTGCTGT 335
 99 lThrGluAlaArgAspArgAlaGlyIleAsnIleThrThrValGluArgA 116
 |||
 336 CACGGAGGCCCGAGACCCGCTGCGGCAACATCACCGATGGAGAGGG 385
 116 spGlyTyrLeuTrrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
 |||
 386 ACGAATACCTCTGGGAAGAGCCCAACAGGTTCCACGCTTGTGATGCA 435
 133 MetIleuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuG 149
 |||
 436 ATGCTCACCATGTGTGGACACAGTGTAAAGGATGCGCTTGTGGG 485
 149 YAspProAsnAlaProArgPheValLeuTrrpLysGlyLeuArgPrrp 166
 |||
 486 GGATCTGTATGACCTCGTGTGTGTGTGGAACAGGAAATGGAGCCGG 535
 166 aLProSerLysLeuThrAspLeuProPheAspLeuMetSerIlePro 182
 |||
 536 TCCCCGGGAAGCTGACTGATTTGCCCTTTCTTGACTTGTATGAGCATTTG 585
 183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
 |||
 586 GGCAAATTCAGGGCTGCTTGGTCCCTTGGAAATTCGCGCTCCCTCC 635
 199 GLyHsIsgluGluSerValIgluGlnPheValArgArgAsnLeuGlyG 216
 |||
 636 AGGATGATGAGAAATCGTGAAGAGTGTTCCTGCGAACCCTTGGATGG 685
 216 lValAlaPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
 |||
 686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTCAGGGGCTTATGACAGGC 735
 233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrrpLys 249
 |||
 736 GATCCTTCAAAAATTAAGATGAAAGCAACATTCGGAAGATTTGGAGCT 785
 249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
 |||
 786 GGAATAAAATGATGATGATTTATTTGTTGATGAACTTCAAGCAATACAG 835
 266 lUArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282


```

102 TCTTGGCCCTCCCTCCATTCCTCCCAACCTCTTTCTTCCACCTCTCCACTC 151
19 tSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerI 36
152 GAAA.....ATTCC.....T 162
36 Le.SerLysArgAsnSerValAlaAsnCysAsnGlyTrpArgThrArgCysSe 52
163 GCCTCTCCCTCCAACTTAT.....CTAACCTGCTC 194
52 rValAlaLysAspIleValMetValAlaValAspIleGlyIleSerGlyLeu 68
195 CATTTCGGAGAAATCCACCCGCTCTCCGCAAAAACCAAGACATCCCGCC 244
68 rValAlaLysAspIleValMetValAlaValAspIleGlyIleSerGlyLeu 84
245 CC.....GTGACTGCTCTCCGCGGAGGCGGATGAGCGGCTC 285
85 CysIleAlaGlnValMetSerAlaAsnIle.....ProAsnLeuMetVal 99
286 TGCAATTCGCGCCAGCCCTCCGCAACCAACGCAAGTGCACCAACTGCTGCT 335
99 LThrGlnAlaArgAspArgAlaGlyIleAsnIleThrValGlnArg 116
336 CACGGAGCCCGAAGACCGCTGCGGCAACATCACACATGAGAGAGG 385
116 spGlyIleValMetValAlaValAspIleGlyIleSerGlyLeu 132
386 ACGGATTCCTCTGGGAGAGGCGCCCAACGCTTCCAGCTTCTGATCCA 435
133 MetLeuThrMetAlaValAlaAspCysGlyLeuLysAspLeuValLeuG 149
436 ATGCTCCCAAGTGTGTGGACAGTGTAAAGATGACCTTGTGTTGGG 485
149 YAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgPro 166
486 GGATCCGATGCCACCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 535
166 a1ProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
536 TGCCCGGAGAGCTGACTGATGATTTGGCTTCTTGTGACTGATGAGACTGGT 585
183 GlyLysLeuArgAlaGlyIlePheGlyProIleGlyLeuArgProSerPro 199
586 GGCMAAAATCAGGGCTGTGGCTTGTGGTGGATTCGGCTCTCCCTCC 635
199 oGlyHisGlnGluSerValGlnPheValArgArgAsnLeuGlyGly 216
636 AGGCATAGAGGAATCGGTGAAGAGTTTGTTCGTCCGAACTTGTGGATG 685
216 LuValPheGlnValGluLeuIleGlnProPheCysSerGlyValIleGly 232
686 AGGTTTTTGAACGGTGTAGCATTTTGTGGTGAAGTTCAAAAGCAATCAAG 735
233 AspProSerLysLeuSerMetLysAlaIlePheGlyLysValIlePheGly 249
736 GATCCCTTCAAAATTAATGATGAAGCAAGCAATTCGGGAAAGTGTGGAGGCT 785
249 uGlnGluThrGlyGlySerIleIleGlyIleGlyIleGlyIlePheLysAla 266
786 GGAATAAATAAGTGTAGCATTTTGTGGTGAAGTTCAAAAGCAATCAAG 835
266 LuArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
836 AAGGAAATGAGCTTCAAAAACCAACCTCCGAGATCCGGCTCCGCAAAAACA 885
283 LysGlyIleThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
886 AAGGTCACAGCTGTGATCTTCCGGAGGAGCTTACCATGATGCTGCTCA 935
299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316

```

```

936 TGCAATTTCTGCCAGACTAGCAACAAGTAATGATCTGGAGCTTT 985
316 eSerIleThrLysSerGlnLysGlyIleYrHisLeuThrYrGluThr 332
986 CAAGTATTAATAACGAGATGAGAGAGTACGTTGACATGTGAAACA 1035
333 ProGlnGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
1036 CCAGAAAGAGTGGTTCCTTTCGAGTGCAAAACACTGTTCTCCACATTC 1085
349 oSerYrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAla 366
1086 TTCCTATGTTGTAAGTACATTTGCTGGCTCCTGCTGCTGGCTGGCAG 1135
366 spAlaLeuSerAsnPheYrYrProProValGlyValAlaThrIleSer 382
1136 ATCCACTTTCAAAAGTGTATTAACCTCCAGTTCAGTTCACATTC 1185
383 YrProGlnGlnAlaIleArgAspGlnArgLeuValAspGlyIleLeu 399
1186 TATCCAAAAGAGCTTATGATACAAATGCTTGTGATGATGATGATG 1235
399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGly 416
1236 GGGGTTTGGCAATTCATCCACGATGCAAGGAGTGAAGATTAAGGA 1285
416 hTrIleYrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
1286 CTATATAACGCTCATCATCTATTCACCAACCGAGCACACCTGGAGG 1335
433 LeuLeuLeuAsnTrpIleGlyGlyAlaLysAsnProGlnIleLeuSer 449
1336 CTACTCTTGAATTAACAATTGAGAGAGCAACTAATCTGGAATTTTAT 1385
449 sThrGlnSerGlnLeuValGlnValValAspArgAspLeuArgLysMet 466
1386 GACGGACAGTGAACCTGTGGAAACAGTTGATGAGATTTGAAAGAAA 1435
466 euIleYrProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
1436 TATATAACCAAAATGCCAGATTCATTTGTGTGTGTGTGTGTGTGTGT 1485
483 ProGlnAlaIleProGlnPheLeuValGlnHisLeuAspThrLeuSer 499
1486 CCTCAAGCTATTCACAGATTTTGTAGTGGCCATCTGATCTGATGAT 1535
499 rAlaLysAlaIleMetAsnAspAsnGlyLeuGlnGlyLeuPheLeuGly 516
1536 TGCATAAGCTTCTATACAGAAATGAGGTTTGAAGGCTCTTCCCTGG 1585
516 LysnTrpValSerGlyValAlaLeuGlyArgCysValGlnGlyValatyr 532
1586 GTAATTTATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1635
533 GluValAlaSerGlnValThrGlyPheLeuSerArgYrAlaIleYrLys 548
1636 GAGGTAGGAGCTGAAAGTAAACGATTTCTCAAAAATAGAGTGCACAAA 1683

```

```

seq_name: /cgn2_6/ptodata/1/ina/6a_COMB.seq.us-09-102-420B-11
seq_documentation_block:
? Sequence ID, Application US/09102420B
? Patent No. 6084155
? GENERAL INFORMATION:
? APPLICANT: Voltrath, Sandra
? APPLICANT: Johnson, Marie
? APPLICANT: Ward, Eric
? APPLICANT: Heifetz, Peter
? TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
? NUMBER OF SEQUENCES: 43
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: No. 6084155artis Corporation

```


